

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
19 September 2002 (19.09.2002)

PCT

(10) International Publication Number  
**WO 02/072892 A1**

(51) International Patent Classification<sup>7</sup>: **C12Q 1/68**,  
C07H 21/02, 21/04

(21) International Application Number: PCT/US02/08187

(22) International Filing Date: 12 March 2002 (12.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/275,232 12 March 2001 (12.03.2001) US

(71) Applicant (for all designated States except US): **CALIFORNIA INSTITUTE OF TECHNOLOGY** [US/US];  
1200 East California Boulevard, MC 201-85, Pasadena, CA  
91125 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **QUAKE, Stephen**  
[US/US]; 744 Plymouth Road, San Marino, CA 91108  
(US). **BRASLAVSKY, Ido** [IL/US]; 120 South Chester  
Avenue, #7, Pasadena, CA 91106 (US). **HEBERT,**  
**Benedict** [CA/US]; 1028 East Del Mar Boulevard, #306,  
Pasadena, CA 91106 (US). **KARTALOV, Emil** [BG/US];  
MC5027, Avery House, 1200 East California Boulevard,  
Pasadena, CA 91125 (US).

(74) Agents: **WANG, Hugh et al.**; Townsend and Townsend  
and Crew LLP, Two Embarcadero Center, Eighth Floor, San  
Francisco, CA 94111-3834 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AT  
(utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA,  
CH, CN, CO, CR, CU, CZ, CZ (utility model), DE, DE  
(utility model), DK, DK (utility model), DM, DZ, EC, EE,  
EE (utility model), ES, FI, FI (utility model), GB, GD, GE,  
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ,  
LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN,  
MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG,  
SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA,  
UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,  
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent  
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,  
NE, SN, TD, TG).

**Declarations under Rule 4.17:**

- of inventorship (Rule 4.17(iv)) for US only
- of inventorship (Rule 4.17(iv)) for US only
- of inventorship (Rule 4.17(iv)) for US only
- of inventorship (Rule 4.17(iv)) for US only

**Published:**

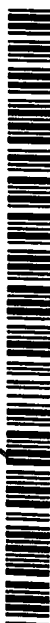
- with international search report
- before the expiration of the time limit for amending the  
claims and to be republished in the event of receipt of  
amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND APPARATUS FOR ANALYZING POLYNUCLEOTIDE SEQUENCES BY ASYNCHRONOUS BASE EXTENSION

(57) Abstract: The invention provides methods and apparatus for analyzing polynucleotide sequences by asynchronous base extension. Some applications of the invention utilize total internal reflection fluorescence microscopy to image polynucleotide molecules at single molecule resolution.

BEST AVAILABLE COPY



WO 02/072892 A1

# **METHODS AND APPARATUS FOR ANALYZING POLYNUCLEOTIDE SEQUENCES BY ASYNCHRONOUS BASE EXTENSION**

## **CROSS-REFERENCES TO RELATED APPLICATIONS**

5           This nonprovisional patent application claims the benefit of U.S. Provisional Patent Application No. 60/275,232, filed March 12, 2001, the disclosure of which is hereby incorporated by reference in its entirety and for all purposes.

## **TECHNICAL FIELD**

10           The present invention relates to novel methods and apparatus for analyzing polynucleotide sequences with high sensitivity and parallelism.

## **BACKGROUND OF THE INVENTION**

15           Methods for analyzing polynucleotide sequences can be grouped to two major fields: electrophoretic and non-electrophoretic methods. The electrophoretic methods include slab gel electrophoresis, capillary electrophoresis, microfabricated capillary arrays, and free solution electrophoresis. All these methods rely on the Sanger method in which polynucleotide chain elongation inhibitors are incorporated into the polynucleotide strands which are then separated according to their sizes, usually on a polyacrylamide gel. These  
20           methods are the common means for analyzing polynucleotide sequences nowadays. However, the process is time-consuming, requires large amount of target polynucleotides and reaction reagents, and has limited ability to read long sequences that are inherent in the gel electrophoresis methods. The non-electrophoretic methods include pyrosequencing, sequencing by hybridization, massively parallel signature sequencing, and sequencing by  
25           mass spectrometry. These methods also have a number of disadvantages. For example, they usually require synchronization of the polynucleotide templates which inevitably decay with each cycle of sequencing reaction.

30           Thus, there is a need in the art for better methods for analyzing polynucleotide sequences, e.g., methods with high throughput, parallelism, and resolution. The present invention fulfills this and other needs.

## SUMMARY OF THE INVENTION

In one aspect, the present invention provides methods for analyzing the sequence of a target polynucleotide. The methods include the steps of (a) providing a primed target polynucleotide immobilized to a surface of a substrate; wherein the target  
5 polynucleotide is attached to the surface with single molecule resolution; (b) In the presence of a polymerase, adding a first fluorescently labeled nucleotide to the surface of the substrate under conditions whereby the first nucleotide attaches to the primer, if a complementary nucleotide is present to serve as template in the target polynucleotide; (c) determining presence or absence of a fluorescence signal on the surface where the target polynucleotide is  
10 immobilized, the presence of a signal indicating that the first nucleotide was incorporated into the primer, and hence the identity of the complementary base that served as a template in the target polynucleotide; and (d) repeating steps (b)-(c) with a further fluorescently labeled nucleotide, the same or different from the first nucleotide, whereby the further nucleotide attaches to the primer or a nucleotide previously incorporated into the primer.

15 In some methods, a plurality of different primed target polynucleotides are immobilized to different portions of the substrate. In some methods, steps (b)-(c) are performed at least four times with four different types of labeled nucleotides. In some methods, steps (b)-(c) are performed until the identity of each base in the target polynucleotide has been identified. In some methods, there is an additional step of removing  
20 the signal after step (c). In some methods, all ingredients are present simultaneously and a continuous monitoring of the incorporation is facilitated.

In some methods of the invention, the presence or absence of a fluorescence signal is determined with total internal reflection fluorescence (TIRF) microscopy. In some methods, the target polynucleotide is primed with a fluorescently labeled primer (e.g., with  
25 Cy5 or Cy3). Some methods of the invention employ nucleotides that are labeled with Cy3 or Cy5.

Various materials can be used to immobilize the target polynucleotides. In some methods, a fused silica or glass slide is used. In some methods, the substrate surface is coated with a polyelectrolyte multilayer (PEM). The PEM can be terminated with a  
30 polyanion, which helps to repel nucleotides from the surface and reduce non-specific binding to the surface. The polyanion can bear pendant carboxylic acid groups. In some of these methods, the target polynucleotide is biotinylated, and the substrate surface is coated with streptavidin. Often the surface is coated with biotin prior to coating with streptavidin. In

some methods, the surface is coated with a polyelectrolyte multilayer (PEM) terminated with carboxylic acid groups prior to attachment of biotin.

In some methods of the invention, a light source for illuminating the surface of said substrate and a detection system for detecting a signal from said surface are employed.

5 Optionally, an appropriately programmed computer is also employed for recording identity of a nucleotide when the nucleotide becomes incorporated into the immobilized primer.

In another aspect, the invention provides apparatus for carrying out the methods of the invention. Typically, the apparatus contain (a) a flow cell which houses a substrate for immobilizing target polynucleotide(s) with single molecule resolution; (b) an  
10 inlet port and an outlet port in fluid communication with the flow cell for flowing fluids into and through the flow cell; (c) a light source for illuminating the surface of the substrate; and (d) a detection system for detecting a signal from said surface. Some of the apparatus are microfabricated. In some of these apparatus, the substrate is a microfabricated synthesis channel.

15 A further understanding of the nature and advantages of the present invention may be realized by reference to the remaining portions of the specification, the figures and claims.

All publications, patents, and patent applications cited herein are hereby expressly incorporated by reference in their entirety and for all purposes to the same extent as  
20 if each was so individually denoted.

### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows schematically immobilization of a primed polynucleotide and  
25 incorporation of labeled nucleotides.

Figure 2 shows schematically the optical setup of a detection system for total internal reflection microscopy.

30 Figure 3 shows results which indicate that streptavidin is required for immobilizing the polynucleotide template in an exemplified embodiment.

Figure 4 shows results which indicate that DNA polymerase incorporating labeled nucleotide into the immobilized primer is visualized with single molecule resolution.

Figure 5 shows incorporation of multiple labeled nucleotides in a bulk experiment in solution, using biotin-labeled 7G oligonucleotide template (SEQ ID NO:1) and p7G primer (SEQ ID NO:2).

5

Figure 6 shows low background signal from free nucleotides in solution and detection of signals from incorporated nucleotides.

Figure 7 shows results from experiments and simulation of multiple bleaching.

10

Figure 8 shows dynamics of incorporation of labeled nucleotides into the immobilized primer.

Figure 9 shows multiple incorporation events of labeled nucleotides over a period of time.

15

Figure 10 shows statistics of incorporation of labeled nucleotides over a period of time.

Figure 11 shows correlation between location of labeled primer and location of incorporation of labeled nucleotides.

20

Figure 12 shows correlation graphs for incorporation of two labeled nucleotides, using a 6TA6GC oligonucleotide template (SEQ ID NO:6) and a p7G primer (SEQ ID NO:2). Partial sequences of the template, 5'- GccccccAttttt - 3' (SEQ ID NO:7), and the extended product, 5' - aaaaaaUggggggC (SEQ ID NO:8), are also shown in the Figure.

25

Figure 13 shows detection of fluorescence resonance energy transfer (FRET) when two different labels are incorporated into the same primer. The polynucleotide template used here is the 7G7A oligonucleotide (SEQ ID NO:5), but only part of the sequence, 5' - AttctttGcttcttAttctttGcttcttAttctttG - 3' (SEQ ID NO:9), is shown in the Figure.

30

Figure 14 shows correlation of single molecule FRET signals over a period of time.

Figure 15 shows the expected signals from an experiment in which two colors, donor and acceptor, are incorporated one after the another. Partial sequences of the template, 5'- GcccccAttttt - 3' (SEQ ID NO:7), and the extended product, 5' - aaaaaaUggggggC (SEQ ID NO:8), are also shown in the Figure.

## DETAILED DESCRIPTION

### I. Overview

The present invention provides methods and apparatus for analyzing polynucleotides with high sensitivity, parallelism, and long read frames. The invention is predicated in part on visualization of incorporation of labeled nucleotides into immobilized polynucleotide template molecules in a time resolved manner with single molecule resolution. As each of the immobilized template molecules is read individually, no synchronization is needed between the different molecules. Instead, with methods of the present invention, asynchronous base extension is sufficient for analyzing a target polynucleotide sequence.

In some aspects of the invention, single molecule resolution was achieved by immobilizing the template molecules at very low concentration to a surface of a substrate, coating the surface to create surface chemistry that facilitates template attachment and reduces background noise, and imaging nucleotide incorporation with total internal reflection fluorescence microscopy. Analysis with single molecule resolution provides the advantage of monitoring the individual properties of different molecules. It allows identification of properties of an individual molecule that can not be revealed by bulk measurements in which a large number of molecules are measured together. Furthermore, to determine kinetics, bulk measurements require synchronization of the molecules or system state, while in single molecule analysis there is no need for synchronization.

The polynucleotides suitable for analysis with the invention can be DNA or RNA. The analysis can be for sequence analysis, DNA fingerprinting, polymorphism identification, or gene expression measurement. The methods can also be used to analyze activities of other biomacromolecules such as RNA translation and protein assembly. In a preferred embodiment, the method entails immobilization of primed polynucleotide templates to the surface of a solid substrate (e.g., a glass slide). The templates are pre-hybridized to a

labeled primer (e.g., with a fluorescent dye) so that their location on the surface can be imaged with single molecule sensitivity. An evanescent light field is set up at the surface in order to image the fluorescently labeled polynucleotide molecules. The evanescent field is also used to image fluorescently labeled nucleotide triphosphates (dNTPs or NTPs) upon their incorporation into the immobilized primer when a polymerase is present.

Methods of the present invention find various applications in polynucleotide sequence analysis. In some applications, a static approach is employed. Such an approach involves adding just one type of labeled nucleotide to the extension reaction at any given time. The signal is incorporated into the primer if the next template residue in the target polynucleotide is the complementary type. Otherwise, a different type of labeled nucleotide is used until the correct residue is incorporated. In other applications, a dynamic approach is employed. In these methods, all four types of nucleotides (at least one type labeled) are simultaneously present in the reaction, and incorporation of the signals into the primer is monitored dynamically. For example, incorporated signals are imaged continuously, preferably at a rate faster than the rate at which the nucleotides are incorporated into the primer.

Preferably, visualization of the templates or incorporated nucleotides are realized with total internal reflection (TIR) fluorescence microscopy. With TIR technology, the excitation light (e.g., a laser beam) illuminates only a small volume of liquid close to the substrate (excitation zone). Signals from free nucleotides in solution that are not present in the excitation zone are not detected. Signals from free nucleotides that diffuse into the excitation zone appear as a broad band background because the free nucleotides move quickly across the excitation zone. Optionally, the fluorescence signals are removed by photobleaching or by chemical means after one or more rounds of incorporation. The methods can also employ microfluidic means to control flow of reaction reagents. In such methods, labeled nucleotides and other reaction reagents can be exchanged in a fast and economic way.

Further, employing a microfluidic device which allows fast fluid exchange, concentrations of nucleotides and/or other reaction reagents can be alternated at different time points of the analysis. This could lead to increased incorporation rate and sensitivity of the analysis. For example, when all four types of nucleotides are simultaneously present in the reaction to monitor dynamic incorporation of nucleotides, concentrations of the nucleotides can be alternated between  $\mu\text{M}$  range and sub-nM range. This leads to both better visualization of the signals when low concentrations of nucleotides are present, and increased

polymerization rate when higher concentrations of nucleotides are present. Using a microfluidic device, the rate at which the concentrations can be alternated can be as high as a few tens of Hertz. Alternating concentrations of nucleotides is also beneficial to improving signal visualization and polymerization rate in the static approach of sequence analysis. In this approach, after adding a given type of labeled nucleotide to the immobilized template/primer complex and sufficient time for incorporation, free nucleotides (as well as other reaction reagents in solution) can be flown out using a microfluidic device. This will leave a much lower concentration of free nucleotide when the signals are visualized. Optionally, an additional washing step can be employed to further reduce the free nucleotide concentration before the signals are imaged.

In some methods, polynucleotide sequence analysis is accomplished by using four different fluorescent labels on the four nucleotide triphosphates. Incorporated signals are imaged and then photobleached before the next incorporation cycle. Runs of identical bases (e.g., AAAAA) can be identified by, e.g., monitoring the intensity of the signal so that the number of fluorophores at the emitting spot can be determined. Further, signals due to fluorescence resonance energy transfer (FRET) can be detected from individual DNA strands when two different type of fluorescent dyes are incorporated into the same DNA. Such signals are useful to determine sequence information of the immobilized template polynucleotide.

Thus, in some methods, multiple types of labeled nucleotides (e.g., 2 to 4 types each labeled with a different fluorescent dye) can be added at the same time for the extension reactions. In some methods, one type of labeled nucleotide is added at a step, and each extension cycle may comprise four such steps in order to observe the incorporation of a complementary nucleotide. In some methods, less than all four dNTPs are labeled. For example, the analysis can have only two of the nucleotides labeled. By repeating the experiment with different pairs (e.g., AT, AG, AC, TG, TC, GC), the original nucleotide sequence can be delineated. In some methods, the incorporation/extension reaction is performed with multiple copies of the template polynucleotide. Alternatively, one immobilized template molecule can be used repeatedly, by denaturing the extended molecule, removing the newly synthesized strand, annealing a new primer, and then repeating the experiment in situ with fresh reagents.

The present invention is also useful to obtain partial sequence information of a target polynucleotide, e.g., by using only two or three labeled nucleotide species. The relative positions of two or three nucleotide species in the sequence in conjunction with



known sequence databases can facilitate determination of the identity of the target sequence, i.e., whether it is identical or related to a known sequence. Such an approach is useful, for example, in determining gene expressions by sequencing cDNA libraries.

The present methods avoid many of the problems observed with the prior art sequencing methods. For example, the methods are highly parallel since many molecules are analyzed simultaneously and in high density (e.g., one template molecule per  $\sim 10\mu\text{m}^2$  of surface area). Thus, many different polynucleotides can be sequenced or genotyped on a single substrate surface simultaneously. In addition, stepwise addition of nucleotides is unnecessary in some methods, as all four nucleotides can be added simultaneously. Rather, sequence information is produced continuously as polymerases continually incorporate all four nucleotides into growing polynucleotide chains. The methods are also extremely sensitive because information obtained from only a single copy of the template molecule is needed in order to determine its sequence. Releasing the extension product from the polynucleotide template, e.g., by denaturing and annealing the template with a different primer provides the opportunity to read again the same template molecule with different sets of nucleotides (e.g., different combinations of two types of labeled nucleotide and two types of unlabeled nucleotides).

## II. Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which this invention pertains. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton *et al.*, DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY (2d ed. 1994); THE CAMBRIDGE DICTIONARY OF SCIENCE AND TECHNOLOGY (Walker ed., 1988); and Hale & Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY (1991). Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described. The following definitions are provided to assist the reader in the practice of the invention.

"Array" refers to a solid support having more than one site or location having either a target polynucleotide or a polymerase bound thereto.

A "base" or "base-type" refers to a particular type of nucleoside base. Typical bases include adenine, cytosine, guanine, uracil, or thymine bases where the type refers to the subpopulation of nucleotides having that base within a population of nucleotide triphosphates

bearing different bases. Other rarer bases or analogs can be substituted such as xanthine or hypoxanthine or methylated cytosine.

“Complements a region of the target nucleic acid downstream of the region to be sequenced” in the context of sequencing or genotyping refers to the fact that the primers are extended in a 3' direction by a polymerase. Therefore the primer binds to a subsequence of the target 3' (downstream) to the target sequence that is to be determined as the 3' end of the primer is extended.

“Genotyping” is a determination of allelic content of a target polynucleotide without necessarily determining the sequence content of the entire polynucleotide. It is a subset of sequencing. For example the identification of single nucleotide polymorphisms by determination of single base differences between two known forms of an allele is a form of sequencing that does not require all the target polynucleotide to be sequenced.

“Immobilizing” refers to the attachment of a target nucleic acid or polymerase to a solid support by a means that prevents its release in a reaction solution. The means can be covalent bonding or ionic bonding or hydrophobic bonding.

“Nucleoside” includes natural nucleosides, including ribonucleosides and 2'-deoxyribonucleosides, as well as nucleoside analogs having modified bases or sugar backbones.

The terms “nucleic acid” or “nucleic acid molecule” refer to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, can encompass known analogs of natural nucleotides that can function in a similar manner as naturally occurring nucleotides. Unless otherwise noted, “nucleic acid” and “polynucleotide” are used interchangeably.

“Oligonucleotide” or “polynucleotide” refers to a molecule comprised of a plurality of deoxyribonucleotides or nucleoside subunits. The linkage between the nucleoside subunits can be provided by phosphates, phosphonates, phosphoramidates, phosphorothioates, or the like, or by nonphosphate groups as are known in the art, such as peptide-type linkages utilized in peptide nucleic acids (PNAs). The linking groups can be chiral or achiral. The oligonucleotides or polynucleotides can range in length from 2 nucleoside subunits to hundreds or thousands of nucleoside subunits. While oligonucleotides are preferably 5 to 100 subunits in length, and more preferably, 5 to 60 subunits in length, the length of polynucleotides can be much greater (e.g., up to 100 kb). (...if a whole chromosome is targeted...Thought 100kb will be already nice..) [“e.g.” means it is not exclusive. Also, “100 Mb” probably does not make practical sense]

"Optical reader" or "detection system" refers to a device that can detect and record light emitted from the labeled dNTP (or NTP) or immobilized polynucleotide template (and/or primer) molecules.

The term "primer" refers to an oligonucleotide, whether occurring naturally as  
5 in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, (i.e., in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature, buffer and pH). The primer is preferably single stranded for maximum  
10 efficiency in amplification, but can alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. Preferably, the primer is an oligodeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of the primers depend on many factors, including temperature, source of primer and  
15 the use of the method.

A primer is selected to be "substantially" complementary to a strand of specific sequence of the template. A primer must be sufficiently complementary to hybridize with a template strand for primer elongation to occur. A primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment can  
20 be attached to the 5' end of the primer, with the remainder of the primer sequence being substantially complementary to the strand. Non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient  
complementarity with the sequence of the template to hybridize and thereby form a template primer complex for synthesis of the extension product of the primer. The use of random  
25 primer is used in some cases. For example, when the terminal sequence of the target or template polynucleotide is not known, random primer combinations can be used.

The term "probe" refers to an oligonucleotide (i.e., a sequence of nucleotides), whether occurring naturally as in a purified restriction digest or produced synthetically, recombinantly or by PCR amplification, which is capable of hybridizing to another  
30 oligonucleotide of interest. A probe can be single-stranded or double-stranded. Probes are useful in the detection, identification and isolation of particular gene sequences. It is contemplated that any probe used in the present invention can be labeled with any "reporter molecule," so that is detectable in any detection system, including, but not limited to fluorescent, enzyme (e.g., ELISA, as well as enzyme-based histochemical assays),

radioactive, quantum dots, and luminescent systems. It is not intended that the present invention be limited to any particular detection system or label.

"Sequencing" refers to the determination of the order and position of bases in a polynucleotide molecule.

5 "Single molecule configuration" refers to an array of molecules on a solid support where members of the array are present as an individual molecule located in a defined location. The members can be the same or different.

"Single molecule resolution" refers to the ability of a system to resolve one molecule from another. For example, in far field optical system the detection limit is in the  
10 order of a micron. This implies that the distance between two identical molecules to be resolved is at least few microns apart.

"Specific hybridization" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions. Stringent conditions are conditions under which a probe can hybridize to its target subsequence, but to  
15 no other sequences. Stringent conditions are sequence-dependent and are different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5° C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50%  
20 of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Typically, stringent conditions include a salt concentration of at least about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides). Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide or tetraalkyl ammonium  
25 salts. For example, conditions of 5X SSPE (750 mM NaCl, 50 mM Na Phosphate, 5 mM EDTA, pH 7.4) and a temperature of 25-30°C are suitable for allele-specific probe hybridizations. (*See Sambrook et al., Molecular Cloning 2001*).

The term "template" or "target" refers to a polynucleotide of which the sequence is to be analyzed. In some cases "template" is sought to be sorted out from other  
30 polynucleotide sequences. "Substantially single-stranded template" is polynucleotide that is either completely single-stranded (having no double-stranded areas) or single-stranded except for a proportionately small area of double-stranded polynucleotide (such as the area defined by a hybridized primer or the area defined by intramolecular bonding). "Substantially double-stranded template" is polynucleotide that is either completely double-stranded (having

no single-stranded region) or double-stranded except for a proportionately small area of single-stranded polynucleotide.

### III. Template Preparation and Immobilization

#### 5 A. Introduction

This invention provides novel methods and apparatus to analyze polynucleotide sequences (e.g., sequencing and genotyping). Preferably, the target or template polynucleotide to be analyzed is immobilized to the surface of a solid substrate (e.g., a fused silica slide) at single molecule resolution. Preferably, the polynucleotide is pre-  
10 hybridized to a labeled primer. A DNA or RNA polymerase, four different types of nucleotide triphosphates (NTPs or dNTPs, depending on the template and polymerase used), and other reaction reagents are then applied to the immobilized polynucleotide. At least one type of the nucleotides are fluorescently labeled. When more than one type of NTPs are labeled, the labels are preferably different for different NTPs. Using TIR fluorescent  
15 microscopy, incorporation of the labeled nucleotide into a target or template polynucleotide is detected by imaging fluorescence signal from the immobilized polynucleotide with single molecule resolution. Preferably, all four labeled NTPs are present simultaneously. As the polymerase continues to move along the target polynucleotide, the polynucleotide sequence is read from the order of the incorporated labels.

#### 20 B. Target or template polynucleotide

The target polynucleotide is not critical and can come from a variety of standard sources. It can be mRNA, ribosomal RNA, genomic DNA or cDNA. They can comprise naturally occurring and or non-naturally occurring nucleotides. Templates suitable for analysis according to the present invention can have various sizes. For example, the  
25 template can have a length of 100 bp, 200 bp, 500 bp, 1 kb, 3 kb, 10 kb, or 20 kb and so on. When the target is from a biological source, there are a variety of known procedures for extracting polynucleotide and optionally amplified to a concentration convenient for genotyping or sequence work. Polynucleotide can be obtained from any living cell of a person, animal or plant. Humans, pathogenic microbes and viruses are particularly  
30 interesting sources.

Polynucleotide amplification methods are known in the art. Preferably, the amplification is carried out by polymerase chain reaction (PCR). See, U.S. Pat. Nos. 4,683,202, 4,683,195 and 4,889,818; Gyllenstein et al., 1988, Proc. Natl. Acad. Sci. USA 85:

7652-7656; Ochman et al., 1988, Genetics 120: 621-623; Loh et al., 1989, Science 243: 217-220; Innis et al., 1990, PCR Protocols, Academic Press, Inc., San Diego, Calif. Other amplification methods known in the art that can be used in the present invention include ligase chain reaction (see EP 320,308), or methods disclosed in Kricka et al., 1995,  
5 *Molecular Probing, Blotting, and Sequencing*, Chap. 1 and Table IX, Academic Press, New York.

C. Primer annealing

Primers in combination with polymerases are used to sequence target  
10 polynucleotide. Primer length is selected to provide for hybridization to complementary template polynucleotide. The primers will generally be at least 10 bp in length, usually between 15 and 30 bp in length. If part of the template sequence is known, a specific primer can be constructed and hybridized to the template. Alternatively, if sequence of the template is completely unknown, the primers can bind to synthetic oligonucleotide adaptors joined to  
15 the ends of target polynucleotide by a ligase.

In some methods, the primer is labeled. When hybridized to the immobilized template, the labeled primer facilitates imaging location of the template. As exemplified in the Examples below, the primer can be labeled with a fluorescent label (e.g., Cy5). Preferably, the label used to label the primer is different from the labels on the nucleotides in  
20 the subsequent extension reactions.

The primers can be synthetically made using conventional nucleic acid synthesis technology. For example, the primers can be conveniently synthesized on an automated DNA synthesizer, e.g. an Applied Biosystems, Inc. (Foster City, Calif.) model 392 or 394 DNA/RNA Synthesizer, using standard chemistries, such as phosphoramidite  
25 chemistry, e.g. disclosed in the following references: Beaucage and Iyer, Tetrahedron, 48: 2223-2311 (1992); Molko et al, U.S. Pat. No. 4,980,460; Koster et al, U.S. Pat. No. 4,725,677; Caruthers et al, U.S. Pat. Nos. 4,415,732; 4,458,066; and 4,973,679; and the like. Alternative chemistries, e.g. resulting in non-natural backbone groups, such as phosphorothioate, phosphoramidate, and the like, may also be employed provided that the  
30 resulting oligonucleotides are compatible with the polymerase. The primers can also be ordered commercially from a variety of companies which specialize in custom oligonucleotides such as Operon Inc (Alameda, California).

Primer annealing is performed under conditions which are stringent enough to achieve sequence specificity yet sufficiently permissive to allow formation of stable hybrids

at an acceptable rate. The temperature and length of time required for primer annealing depend upon several factors including the base composition, length and concentration of the primer, and the nature of the solvent used, e.g., the concentration of DMSO, formamide, or glycerol, and counter ions such as magnesium. Typically, hybridization with synthetic polynucleotides is carried out at a temperature that is approximately 5 to 10°C below the melting temperature of the target-primer hybrid in the annealing solvent. In some methods, the annealing temperature is in the range of 55 to 75°C. and the primer concentration is approximately 0.2 µM. Other conditions of primer annealing are provided in the Examples below. Under these preferred conditions, the annealing reaction can be complete in only a few seconds.

#### D. Immobilization of template polynucleotide

Preferably, the template or target polynucleotide molecules are provided as single molecule arrays immobilized to the surface of a solid substrate. The substrate can be glass, silica, plastic or any other conventionally non-reactive material that will not create significant noise or background for the fluorescent detection methods. Substrate surface to which the template polynucleotides are to be immobilized can also be the internal surface of a flow cell in a microfluidic apparatus, e.g., a microfabricated synthesis channel of the apparatus as described in the PCT application of Quake et al. (WO 01/32930; which is incorporated herein by reference). In some preferred embodiments, the solid support is made from fused silica slide (e.g., a fused silica glass slide from Esco, Cat. R130110). Compared to other support materials (e.g., a regular glass slide), fused silica has very low auto-fluorescence.

In some applications of the present invention, the template or target polynucleotides are immobilized to the substrate surface with single molecule resolution. In such methods, as exemplified in the Examples below, single molecule resolution is achieved by using very low concentration of the polynucleotide in the immobilization reaction. For example, a 10 pM concentration for a 80-mer polynucleotide template allows attachment of the polynucleotide to the surface of a silica slide at single molecule resolution (see Example 1). Template immobilization with single molecule resolution can also be verified by measuring bleach pattern of the fluorescently labeled templates (see Example 5).

In some methods, the templates are hybridized to the primers first and then immobilized to the surface. In some methods, the templates are immobilized to the surface prior to hybridization to the primer. In still some methods, the primers are immobilized to the

surface, and the templates are attached to the substrates through hybridization to the primers. In still some methods, the polymerase is immobilized to the surface.

Various methods can be used to immobilize the templates or the primers to the surface of the substrate. The immobilization can be achieved through direct or indirect bonding of the templates to the surface. The bonding can be by covalent linkage. See, Joos et al., *Analytical Biochemistry* 247:96-101, 1997; Oroskar et al., *Clin. Chem* 42:1547-1555, 1996; and Khandjian, *Mole. Bio. Rep.* 11:107-115, 1986. The bonding can also be through non-covalent linkage. For example, Biotin-streptavidin (Taylor et al., *J. Phys. D. Appl. Phys.* 24:1443, 1991) and digoxigenin and anti-digoxigenin (Smith et al., *Science* 253: 1122, 1992) are common tools for attaching polynucleotides to surfaces and parallels. Alternatively, the bonding can be achieved by anchoring a hydrophobic chain into a lipidic monolayer or bilayer. When biotin-streptavidin linkage is used to immobilize the templates, the templates are biotinylated, and one surface of the substrates are coated with streptavidin. Since streptavidin is a tetramer, it has four biotin binding sites per molecule. Thus, it can provide linkage between the surface and the template. In order to coat a surface with streptavidin, the surface can be biotinylated first, and then parts of the four binding sites of streptavidin can be used to anchor the protein to the surface, leaving the other sites free to bind the biotinylated template (see, Taylor et al., *J. Phys. D. Appl. Phys.* 24:1443, 1991). Such treatment leads to a high density of streptavidin on the surface of the substrate, allowing a correspondingly high density of template coverage. Surface density of the template molecules can be controlled by adjusting concentration of the template which is applied to the surface. Reagents for biotinylating a surface can be obtained, for example, from Vector laboratories. Alternatively, biotinylation can be performed with BLCPA: EZ-Link Biotin LC-PEO-Amine (Pierce, Cat. 21347).

In some methods, labeled streptavidin (e.g., with a fluorescent label) of very low concentration (e.g., in the  $\mu\text{M}$ , nM or pM range) is used to coat the substrate surface prior to template immobilization. This facilitates immobilization of the template with single molecule resolution. It also allows monitoring of spots on the substrate to which the template molecules are attached, and subsequent nucleotide incorporation events.

While diverse polynucleotide templates can be each immobilized to and sequenced in a separate substrate, multiple templates can also be analyzed on a single substrate. In the latter scenario, the templates are attached at different locations on the substrate. This can be accomplished by a variety of different methods, including hybridization of primer capture sequences to oligonucleotides immobilized at different points



on the substrate, and sequential activation of different points down the substrate towards template immobilization.

Methods of creation of surfaces with arrays of oligonucleotides have been described, e.g., in U.S. Patent Nos. 5,744,305, 5,837,832, and 6,077,674. Primers with two domains, a priming domain and a capture domain, can be used to anchor templates to the substrate. The priming domain is complementary to the target template. The capture domain is present on the non-extended side of the priming sequence. It is not complementary to the target template, but rather to a specific oligonucleotide sequence present on the substrate. The target templates can be separately hybridized with their primers, or (if the priming sequences are different) simultaneously hybridized in the same solution. Incubation of the primer/template duplexes with the substrate under hybridization conditions allows attachment of each template to a unique spot. Multiple substrates can be charged with templates in this fashion simultaneously.

Another method for attaching multiple templates to the surface of a single substrate is to sequentially activate portions of the substrate and attach template to them. Activation of the substrate can be achieved by either optical or electrical means. Optical illumination can be used to initiate a photochemical deprotection reaction that allows attachment of the template to the surface (see, e.g., U.S. Patent Nos. 5,599,695, 5,831,070, and 5,959,837). For instance, the substrate surface can be derivitized with "caged biotin", a commercially available derivative of biotin that becomes capable of binding to avidin only after being exposed to light. Templates can then be attached by exposure of a site to light, filling the channel with avidin solution, washing, and then flowing biotinylated template into the channel. Another variation is to prepare avidinylated substrate and a template with a primer with a caged biotin moiety; the template can then be immobilized by flowing into the channel and illumination of the solution above a desired area. Activated template/primer duplexes are then attached to the first wall they diffused to, yielding a diffusion limited spot.

Electrical means can also be used to direct template to specific locations on a substrate. By positively charging one electrode in the channel and negatively charging the others, a field gradient can be created which drives the template to a single electrode, where it can attach (see, e.g., U.S. Patent Nos. 5,632,957, 6,051,380, and 6,071,394). Alternatively, it can be achieved by electrochemically activating regions of the surface and changing the voltage applied to the electrodes. Patterning of particular chemicals, include proteins and DNA is possible with a stamp method, in which a microfabricated plastic stamp is pressed on the surface (see, e.g., Lopez et al., J. Amer. Chem. Soc. 115:10774-81, 1993). Different

templates can also be attached to the surface randomly as the reading of each individual is independent from the others.

E. Treatment of substrate surface

5 In some applications, surface of the substrate is pretreated to create surface chemistry that facilitates attachment of the polynucleotide templates and subsequent synthesis reactions. The surface chemistry also reduces the background from non specific attachment of free labeled nucleotide to the surface of the substrate.

10 In some methods, the surface is coated with a polyelectrolyte multilayer (PEM). In some methods, non-PEM based surface chemistry can be created prior to template attachment. Preferably, the substrate surface is coated with a polyelectrolyte multilayer (PEM). Attachment of templates to PEM-coated surface can be accomplished by light-directed spatial attachment (see, e.g.; U.S. Patent Nos. 5,599,695, 5,831,070, and 5,959,837). Alternatively, the templates can be attached to PEM-coated surface entire chemically (see  
15 below for detail).

PEM formation has been described in Decher et al. (Thin Solid Films, 210:831-835, 1992). PEM formation proceeds by the sequential addition of polycations and polyanions, which are polymers with many positive or negative charges, respectively. Upon addition of a polycation to a negatively-charged surface, the polycation deposits on the  
20 surface, forming a thin polymer layer and reversing the surface charge. Similarly, a polyanion deposited on a positively charged surface forms a thin layer of polymer and leaves a negatively charged surface. Alternating exposure to poly(+) and poly(-) generates a polyelectrolyte multilayer structure with a surface charge determined by the last  
polyelectrolyte added; in the case of incompletely-charged surfaces, multiple-layer deposition  
25 also tends to increase surface charge to a well defined and stable level.

An exemplified scheme of coating a substrate with PEM for immobilizing polynucleotide is provided in PCT publication WO 01/32930. Detailed procedures are also disclosed in the Examples below. Briefly, the surface of the substrate (e.g., a glass cover slip) is cleaned with a RCA solution. After cleaning, the substrate is coated with a  
30 polyelectrolyte multilayer (PEM). Following biotinylation of the carboxylic acid groups, streptavidin is then applied to generate a surface capable of capturing biotinylated molecules. Biotinylated polynucleotide templates are then added to the coated glass cover slip for attachment. The surface chemistry thus created provides various advantages for the methods of the present invention, because it generates a strong negatively-charged surface which

repels the negatively-charged nucleotides. First, a polyelectrolyte multilayer terminated with carboxylic acid-bearing polymer is easy to attach polynucleotide to because carboxylic acids are good targets for covalent bond formation. In addition, the attached template is active for extension by polymerases – most probably, the repulsion of like charges prevents the template from “laying down” on the surface. Finally, the negative charge repels the fluorescent nucleotides, and nonspecific binding is low.

The attachment scheme described here is easy to generalize on. Without modification, the PEM/biotin/streptavidin surface that is produced can be used to capture or immobilize any biotinylated molecule. A slight modification can be the use of another capture pair, e.g., substituting digoxigenin (dig) for biotin and labeling the molecule to be immobilized with anti-digoxigenin (anti-dig). Reagents for biotinylation or dig-labeling of amines are all commercially available.

Another generalization is that the chemistry is nearly independent of the surface chemistry of the support. Glass, for instance, can support PEMs terminated with either positive or negative polymer, and a wide variety of chemistry for either. But other substrates such as silicone, polystyrene, polycarbonate, etc, which are not as strongly charged as glass, can still support PEMs. The charge of the final layer of PEMs on weakly-charged surfaces becomes as high as that of PEMs on strongly-charged surfaces, as long as the PEM has sufficiently-many layers. This means that all the advantages of the glass/PEM/biotin/Streptavidin/biotin-DNA surface chemistry can be applied to other substrates.

#### IV. Primer Extension Reaction

Once templates are immobilized to the surface of a substrate, primer extension reactions are performed, e.g., as described in Sambrook, *supra*; Ausubel, *supra*; and Hyman, Anal. Biochem., 174, p. 423, 1988. In some methods, the primer is extended by a polynucleotide polymerase in the presence of a single type of labeled nucleotide. In other methods, all four types of differently labeled nucleotides are present. In some applications of the present invention, a combination of labeled and non-labeled nucleotides are used in the analysis. A label is incorporated into the template/primer complex only if the specific labeled nucleotide added to the reaction is complementary to the nucleotide on the template adjacent the 3' end of the primer. Optionally, the template is subsequently washed to remove any unincorporated label, and the presence of any incorporated label is determined. As some

errors can be caused by the polymerase, the reaction conditions and incubation time should minimize these errors.

A. Labeled nucleotides

To facilitate detection of nucleotide incorporation, at least one and usually all types of the deoxyribonucleotides (dATP, dTTP, dGTP, dCTP, dUTP/dTTP) or nucleotides (ATP, UTP, GTP, and CTP) are labeled with fluorophores. When more than one type of nucleotides are labeled, a different kind of label can be used to label each different type of nucleotide. However, in some applications, the different types of nucleotides can be labeled with the same kind of labels.

Various fluorescent labels can be used to label the nucleotides in the present invention. The fluorescent label can be selected from any of a number of different moieties. The preferred moiety is a fluorescent group for which detection is quite sensitive. The affinity to the surface could be changed between different dyes. Low affinity to the surface is preferred. For example, Cy3 and Cy5 are used to label the primer or nucleotides in some methods of the invention. However, Cy5 has higher affinity to the surface under certain experimental condition than Cy3.

Other factors that need to be considered include stability of the dyes. For example, Cy5 is less stable and tends to bleach faster than Cy3. Such property can be of advantage or disadvantage, depending on the circumstances. In addition, different sizes of the dyes can also affect efficiency of incorporation of labeled nucleotides. Further, length of the linker between the dye and the nucleotide can impact efficiency of the incorporation (see, Zhu and Waggoner, Cytometry 28: 206, 1997).

An exemplary list of fluorophores, with their corresponding absorption/emission wavelength indicated in parenthesis, that can be used in the present invention include Cy3 (550/565), Cy5 (650/664), Cy7 (750/770), Rho123 (507/529), R6G (528/551), BODIPY 576/589 (576/589), BODIPY TR (588/616), Nile Blue (627/660), BODIPY 650/665 (650/665), Sulfo-IRD700 (680/705), NN382 (778/806), Alexa488 (490/520), Tetramethylrhodamine (550/570). and Rodamine X (575/605).

The fluorescently labeled nucleotides can be obtained commercially (e.g., from NEN DuPont, Amersham, or BDL). Alternatively, fluorescently labeled nucleotides can also be produced by various fluorescence-labeling techniques, e.g., as described in Kambara et al. (1988) "Optimization of Parameters in a DNA Sequenator Using Fluorescence Detection," Bio/Technol. 6:816-821; Smith et al. (1985) Nucl. Acids Res, 13:2399-2412; and

Smith et al. (1986) Nature 321:674-679. Acyl fluoride of Cy5 cyanine dye can also be synthesized and labeled as described in U.S. Patent No. 6,342,326.

There is a great deal of practical guidance available in the literature for providing an exhaustive list of fluorescent and chromogenic molecules and their relevant optical properties (see, for example, Berlman, *Handbook of Fluorescence Spectra of Aromatic Molecules*, 2nd Edition (Academic Press, New York, 1971); Griffiths, *Colour and Constitution of Organic Molecules* (Academic Press, New York, 1976); Bishop, Ed., *Indicators* (Pergamon Press, Oxford, 1972); Haugland, *Handbook of Fluorescent Probes and Research Chemicals* (Molecular Probes, Eugene, 1992) Pringsheim, *Fluorescence and Phosphorescence* (Interscience Publishers, New York, 1949); and the like. Further, there is extensive guidance in the literature for derivatizing fluorophore and quencher molecules for covalent attachment via common reactive groups that can be added to a nucleotide, as exemplified by the following references: Haugland (*supra*); Ullman et al., U.S. Pat. No. 3,996,345; Khanna et al., U.S. Pat. No. 4,351,760.

There are many linking moieties and methodologies for attaching fluorophore moieties to nucleotides, as exemplified by the following references: Eckstein, editor, *Oligonucleotides and Analogues: A Practical Approach* (IRL Press, Oxford, 1991); Zuckerman et al., *Nucleic Acids Research*, 15: 5305-5321 (1987) (3' thiol group on oligonucleotide); Sharma et al., *Nucleic Acids Research*, 19: 3019 (1991) (3' sulfhydryl); Giusti et al., *PCR Methods and Applications*, 2: 223-227 (1993) and Fung et al., U.S. Pat. No. 4,757,141 (5' phosphoamino group via Aminolink<sup>TM</sup>. II available from Applied Biosystems, Foster City, Calif.) Stabinsky, U.S. Pat. No. 4,739,044 (3' aminoalkylphosphoryl group); Agrawal et al., *Tetrahedron Letters*, 31: 1543-1546 (1990) (attachment via phosphoramidate linkages); Sproat et al., *Nucleic Acids Research*, 15: 4837 (1987) (5' mercapto group); Nelson et al., *Nucleic Acids Research*, 17: 7187-7194 (1989) (3' amino group); and the like.

In instances where a multi-labeling scheme is utilized, a wavelength which approximates the mean of the various candidate labels' absorption maxima may be used. Alternatively, multiple excitations may be performed, each using a wavelength corresponding to the absorption maximum of a specific label.

#### B. Other reaction reagents

##### 1. Polymerases

Many polymerases can be selected for use in this invention. Preferred polymerases are able to tolerate labels on the nucleobase. For example, some applications of the present invention employ polymerases that have increased ability to incorporate modified, fluorophore-labeled, nucleotides into polynucleotides. Examples of such polymerases, e.g., mutant bacteriophage T4 DNA polymerases, have been described in U.S. Patent No. 5,945,312.

Depending on the template, either RNA polymerase, DNA polymerases or reverse transcriptase can be used in the primer extension. For analysis of DNA templates, many DNA polymerases are available. Examples of suitable DNA polymerases include, but are not limited to, Sequenase 2.0.RTM., T4 DNA polymerase or the Klenow fragment of DNA polymerase 1, or Vent polymerase. In some methods, polymerases which lack 3' → 5' exonuclease activity can be used (e.g., T7 DNA polymerase (Amersham) or Klenow -exo fragment of DNA polymerase I (New England Biolabs)). In some methods, when it is desired that the polymerase have proof-reading activity, polymerases lacking 3' → 5' exonuclease activity are not used. In some methods, thermostable polymerases such as ThermoSequenase™ (Amersham) or Taquenase™ (ScienTech, St Louis, MO) are used.

In general, the polymerase should have a fidelity (incorporation accuracy) of at least 99% and a processivity (number of nucleotides incorporated before the enzyme dissociates from the DNA) of at least 20 nucleotides, with greater processivity preferred. Examples include T7 DNA polymerase, T5 DNA polymerase, HIV reverse transcriptase, *E. coli* DNA pol I, T4 DNA polymerase, T7 RNA polymerase, Taq DNA polymerase and *E. coli* RNA polymerase, Phi29 DNA polymerase.

The nucleotides used in the methods should be compatible with the selected polymerase. Procedures for selecting suitable nucleotide and polymerase combinations can be adapted from Ruth et al. (1981) *Molecular Pharmacology* 20:415-422; Kutateladze, T., et al. (1984) *Nuc. Acids Res.*, 12:1671-1686; Chidgeavadze, Z., et al. (1985) *FEBS Letters*, 183:275-278.

The polymerase can be stored in a separate reservoir and flowed onto the substrates (or into a flow chamber/cell which houses the substrate) prior to each extension reaction cycle. The enzyme can also be stored together with the other reaction agents (e.g., the nucleotide triphosphates). Alternatively, the polymerase can be immobilized onto the surface of the substrate while the polynucleotide template is added to the solution.

## 2. Blocking agents

In some methods, it may be desirable to employ a chain elongation inhibitor in the primer extension reaction (see, e.g., Dower et al., U.S. Patent No. 5,902,723). Chain elongation inhibitors are nucleotide analogues which either are chain terminators which prevent further addition by the polymerase of nucleotides to the 3' end of the chain by becoming incorporated into the chain themselves. In some methods, the chain elongation inhibitors are dideoxynucleotides. Where the chain elongation inhibitors are incorporated into the growing polynucleotide chain, they should be removed after incorporation of the labeled nucleotide has been detected, in order to allow the sequencing reaction to proceed using different labeled nucleotides. Some 3' to 5' exonucleases, e.g., exonuclease III, are able to remove dideoxynucleotides.

Other than chain elongation inhibitors, a blocking agent or blocking group can be employed on the 3' moiety of the deoxyribose group of the labeled nucleotide to prevent nonspecific incorporation. Optimally, the blocking agent should be removable under mild conditions (e.g., photosensitive, weak acid labile, or weak base labile groups), thereby allowing for further elongation of the primer strand with a next synthetic cycle. If the blocking agent also contains the fluorescent label, the dual blocking and labeling functions are achieved without the need for separate reactions for the separate moieties. For example, the labeled nucleotide can be labeled by attachment of a fluorescent dye group to the 3' moiety of the deoxyribose group, and the label is removed by cleaving the fluorescent dye from the nucleotide to generate a 3' hydroxyl group. The fluorescent dye is preferably linked to the deoxyribose by a linker arm which is easily cleaved by chemical or enzymatic means.

Examples of blocking agents include, among others, light sensitive groups such as 6-nitroveratryloxycarbonyl (NVOC), 2-nitrobenzyloxycarbonyl (NBOC), . $\alpha$ ., $\alpha$ -dimethyl-dimethoxybenzyloxycarbonyl (DDZ), 5-bromo-7-nitroindolinyl, o-hydroxy-2-methyl cinnamoyl, 2-oxymethylene anthraquinone, and t-butyl oxycarbonyl (TBOC). Other blocking reagents are discussed, e.g., in U.S. Ser. No. 07/492,462; Patchornik (1970) J. Amer. Chem. Soc. 92:6333; and Amit et al. (1974) J. Org. Chem. 39:192. Nucleotides

possessing various labels and blocking groups can be readily synthesized. Labeling moieties are attached at appropriate sites on the nucleotide using chemistry and conditions as described, e.g., in Gait (1984) *Oligonucleotide Synthesis: A Practical Approach*, IRL Press, Oxford.

5 C. Reaction conditions

The reaction mixture for the sequencing comprises an aqueous buffer medium which is optimized for the particular polymerase. In general, the buffer includes a source of monovalent ions, a source of divalent cations and a buffering agent. Any convenient source of monovalent ions, such as KCl, K-acetate, NH<sub>4</sub>-acetate, K-glutamate, NH<sub>4</sub>Cl, ammonium sulfate, and the like may be employed, where the amount of monovalent ion source present in  
10 the buffer will typically be present in an amount sufficient to provide for a conductivity in a range from about 500 to 20,000, usually from about 1000 to 10,000, and more usually from about 3,000 to 6,000 micromhos.

The divalent cation may be magnesium, manganese, zinc and the like, where  
15 the cation will typically be magnesium. Any convenient source of magnesium cation may be employed, including MgCl<sub>2</sub>, Mg-acetate, and the like. The amount of Mg ion present in the buffer may range from 0.5 to 20 mM, but will preferably range from about 1 to 12mM, more preferably from 2 to 10mM and will ideally be about 5mM.

Representative buffering agents or salts that may be present in the buffer  
20 include Tris, Tricine, HEPES, MOPS and the like, where the amount of buffering agent will typically range from about 5 to 150 mM, usually from about 10 to 100 mM, and more usually from about 20 to 50 mM, where in certain preferred embodiments the buffering agent will be present in an amount sufficient to provide a pH ranging from about 6.0 to 9.5, where most preferred is pH 7.6 at 25° C. Other agents which may be present in the buffer medium include  
25 chelating agents, such as EDTA, EGTA and the like.

D. Removal of labels and blocking group

By repeating the incorporation and label detection steps until incorporation is detected, the nucleotide on the template adjacent the 3' end of the primer can be identified.  
30 Once this has been achieved, the label should be removed before repeating the process to discover the identity of the next nucleotide. Removal of the label can be effected by removal of the labeled nucleotide using a 3'-5' exonuclease and subsequent replacement with an unlabeled nucleotide. Alternatively, the labeling group can be removed from the nucleotide.



Release of the fluorescence dye can be achieved if a detachable connection between the nucleotide and the fluorescence molecule is used. For example, the use of disulfide bonds enables one to disconnect the dye by applying a reducing agent like dithiothreitol (DTT).

In a further alternative, where the label is a fluorescent label, it is possible to neutralize the label by bleaching it with radiation. Photobleaching can be performed according to methods, e.g., as described in Jacobson et al., "International Workshop on the Application of Fluorescence Photobleaching Techniques to Problems in Cell Biology", Federation Proceedings, 42:72-79, 1973; Okabe et al., J Cell Biol 120:1177-86, 1993; Wedekind et al., J Microsc. 176 Pt 1): 23-33, 1994; and Close et al., Radiat Res 53:349-57, 1973.

If chain terminators or 3' blocking groups have been used, these should be removed before the next cycle can take place. 3' blocking groups can be removed by chemical or enzymatic cleavage of the blocking group from the nucleotide. For example, chain terminators are removed with a 3'-5' exonuclease, e.g., exonuclease III. Once the label and terminators/blocking groups have been removed, the cycle is repeated to discover the identity of the next nucleotide.

#### E. Sample housing.

The solid substrate is optionally housed in a flow chamber having an inlet and outlet to allow for renewal of reactants which flow past the immobilized moieties. The flow chamber can be made of plastic or glass and should either be open or transparent in the plane viewed by the microscope or optical reader. Electro-osmotic flow requires a fixed charge on the solid substrate and a voltage gradient (current) passing between two electrodes placed at opposing ends of the solid support. Pressure driven flow can be facilitated by microfluidic device with an external pressure source or by microfluidic peristaltic pump (see, e.g., Unger et al., Science 288: 113-116, 2000).

The flow chamber can be divided into multiple channels for separate sequencing. Examples of micro flow chambers are described in Fu et al. (*Nat. Biotechnol.* (1999) 17:1109) which describe a microfabricated fluorescence-activated cell sorter with 3 $\mu$ m x 4 $\mu$ m channels that utilizes electro-osmotic flow for sorting. Preferably, the flow chamber contains microfabricated synthesis channels as described in WO01/32930. The polynucleotide templates can be immobilized to the surface of the synthesis channels. These synthesis channels can be in fluid communication with a microfluidic device which controls flow of reaction reagents. Preferred microfluidic devices that can be employed to control flow of reaction reagents in the present invention have been described in WO01/32930.

The present invention also provide apparatus for carrying out the methods of the invention. Other than the substrate to which the target polynucleotides or primers are attached, the apparatus usually comprise a flow chamber in which the substrate is housed. In addition, the apparatus can optionally contain plumbing devices (e.g., an inlet and an outlet port), a light source, and a detection system described herein. Preferably, a microfabricated apparatus as described in WO01/32930 is adapted to house the substrate of the present invention.

V. Detection of Incorporated Signals

A. Detection system in general

Methods for visualizing single molecules of DNA labeled with an intercalating dye include, e.g., fluorescence microscopy as described in Houseal et al., *Biophysical Journal* 56: 507, 1989. While usually signals from a plurality of molecules are to be detected with the sequencing methods of the present invention, fluorescence from single fluorescent dye molecules can also be detected. For example, a number of methods are available for this purpose (see, e.g., Nie et al., *Science* 266: 1013, 1994; Funatsu et al., *Nature* 374: 555, 1995; Mertz et al., *Optics Letters* 20: 2532, 1995; and Unger et al., *Biotechniques* 27:1008, 1999). Even the fluorescent spectrum and lifetime of a single molecule excited-state can be measured (Macklin et al., *Science* 272: 255, 1996). Standard detectors such as a photomultiplier tube or avalanche photodiode can be used. Full field imaging with a two stage image intensified CCD camera can also used (Funatsu et al., supra). Low noise cooled CCD can also be used to detect single fluorescence molecules (see, e.g., Unger et al., *Biotechniques* 27: 1008-1013, 1999; and SenSys spec: <http://www.photomet.com/pdfs/datasheets/sensys/ss1401e.pdf>).

The detection system for the signal or label can also depend upon the label used, which can be defined by the chemistry available. For optical signals, a combination of an optical fiber or charged couple device (CCD) can be used in the detection step. In those circumstances where the matrix is itself transparent to the radiation used, it is possible to have an incident light beam pass through the substrate with the detector located opposite the substrate from the polynucleotides. For electromagnetic labels, various forms of spectroscopy systems can be used. Various physical orientations for the detection system are available and discussion of important design parameters is provided in the art (e.g., Arndt-Jovin et al., *J Cell Biol* 101: 1422-33, 1985; and Marriott et al., *Biophys J* 60: 1374-87, 1991).

Many applications of the invention require the detection of incorporation of fluorescently labeled nucleotides into single template molecules in a solution. The single-molecule fluorescence detection of the present invention can be practiced using optical setups including near-field scanning microscopy, far-field confocal microscopy, wide-field epi-illumination, and total internal reflection fluorescence (TIRF) microscopy. General reviews are available describing this technology, including, e.g., Basche et. al., eds., 1996, Single molecule optical detection, imaging, and spectroscopy, Weinheim:VCM; and Plakhotnik, et. al., Single-molecule spectroscopy, *Ann. Rev. Phys. Chem.* 48: 181-212. In general, the methods involve detection of laser activated fluorescence using microscope equipped with a camera. It is sometimes referred to as a high-efficiency photon detection system (see, e.g., Nie, et. al., 1994, Probing individual molecules with confocal fluorescence microscopy, *Science* 266:1018-1019. Other suitable detection systems are discussed in the Examples below.

Suitable photon detection systems include, but are not limited to, photodiodes and intensified CCD cameras. In a preferred embodiment, an intensified charge couple device (ICCD) camera is used. The use of a ICCD camera to image individual fluorescent dye molecules in a fluid near the surface of the glass slide is advantageous for several reasons. With an ICCD optical setup, it is possible to acquire a sequence of images (movies) of fluorophores. In certain aspects, each of the dNTPs or NTPs employed in the methods has a unique fluorophore associated with it, as such, a four-color instrument can be used having four cameras and four excitation lasers. Preferably the image could be split to four quarters and imaged by a single camera. For example, the micro-imager of Optical Insights LTD is a simple device that splits the image to four different images in four different spectra just in front of the port of the camera. Illumination with only one laser excitation for the four colors is possible if suitable dyes are used (see, e.g., Rosenblum et al, *Nucleic Acids Research* 25:4500, 1997). For example, the BigDyes have single excitation wavelength spectrum and four different emission wavelength spectrums. They can be obtained from Applied Biosystems (see, <http://www.appliedbiosystems.com/products/productdetail.cfm?ID=82>). Nanocrystals are also found to have a variety of emission wavelengths for a given excitation (see, e.g., U.S. Patent No. 6,309,701; and Lacoste et al., *Proc. Natl. Acad. Sci. USA* 97: 9461-6, 2000). Thus, it is possible to use such optical setup to sequence DNA. In addition, many different DNA molecules spread on a solid support (e.g., a microscope slide) can be imaged and sequenced simultaneously.

B. Total internal reflection fluorescence (TIRF) microscopy

In some preferred embodiments, the present invention uses total internal reflection fluorescence (TIRF) microscopy for two-dimensional imaging fluorescence detection. TIRF microscopy is well known in the art. See, e.g., Watkins et al., J Biomed Mater Res 11:915-38, 1977; and Axelrod et al., J Microsc, 129:19-28, 1983. TIRF microscopy uses totally internally reflected excitation light. When a laser beam was totally reflected at the interface between a liquid and a solid substrate (e.g., a glass), the excitation light beam penetrates only a short distance into the liquid. In other words, the optical field does not end abruptly at the reflective interface, but its intensity falls off exponentially with distance. This surface electromagnetic field, called the 'evanescent wave', can selectively excite fluorescent molecules in the liquid near the interface. The thin evanescent optical field at the interface provides low background and enables the detection of single molecules with high signal-to-noise ratio at visible wavelengths (see, M. Tokunaga et al., Biochem. and Biophys. Res. Comm. 235, 47 (1997) and P. Ambrose, Cytometry, 36, 244 (1999)).

TIRF microscopy has been used to examine various molecular or cellular activities, e.g., cell/substrate contact regions of primary cultured rat myotubes with acetylcholine receptors labeled by fluorescent alpha-bungarotoxin, and human skin fibroblasts labeled with a membrane-incorporated fluorescent lipid (see, e.g., Thompson et al., Biophys J. 33:435-54, 1981; Axelrod, J. Cell. Biol. 89: 141-5, 1981; and Burghardt et al., Biochemistry 22:979-85, 1983). TIRF examination of cell/surface contacts dramatically reduces background from surface autofluorescence and debris. TIRF has also been combined with fluorescence photobleaching recovery and correlation spectroscopy to measure the chemical kinetic binding rates and surface diffusion constant of fluorescent labeled serum protein binding (at equilibrium) to a surface (see, e.g., Burghardt et al., Biophys J. 33:455-67, 1981); and Thompson et al., Biophys J, 43:103-14, 1983). Additional examples of TIRF detection of single molecules have been described in Vale et. al., 1996, Direct observation of single kinesin molecules moving along microtubules, Nature 380: 451; and Xu et al., 1997, Direct Measurement of Single-Molecule Diffusion and Photodecomposition in Free Solution, Science 275: 1106-1109.

The penetration of the field beyond the glass depends on the wavelength and the laser beam angle of incidence. Deeper penetrance is obtained for longer wavelengths and for smaller angles to the surface normal within the limit of a critical angle. In typical assays, fluorophores are detected within about 200 nm from the surface which corresponds to the contour length of about 600 base pairs of DNA. In some embodiments, when longer

polynucleotide templates are analyzed, the polymerase rather than the template is immobilized to the surface so the reaction occurs near the surface at all time. In some embodiments, a prism-type TIRF geometry for single-molecule imaging as described by Xu and Yeung is used (*see*, X-H.N. Xu *et al.*, *Science*, **281**, 1650 (1998)). In some embodiments, an objective type TIRF is used to provide space above the objective so that a microfluidic device can be used (*see*, e.g., Tokunaga *et al.*, *Biochem Biophys Res Commun* 235: 47-53, 1997; Ambrose *et al.*, *Cytometry* 36:224;1999; and Braslavsky *et al.*, *Applied Optics* 40:5650, 2001).

Total internal reflection can be utilized with high numerical aperture objectives (ranging between 1.4 and 1.65 in aperture), preferentially using an inverted microscope. The numerical aperture of an objective is a function of the max angle that can be collected (or illuminated) with the objective in a given refractive index of the media (i.e.,  $NA = n \cdot \sin(\theta_{\text{Max}})$ ). If  $\theta_{\text{Max}}$  is larger than  $\theta_{\text{Critic}}$  for reflection, some of the illuminated rays will be totally internal reflected. So using the peripheral of a large NA objective one can illuminate the sample with TIR through the objective and use the same objective to collect the fluorescence light. Therefore, the objective plays double roles as a condenser and an imaging objective.

Single molecule detection can be achieved using flow cytometry where flowing samples are passed through a focused laser with a spatial filter used to define a small volume. US Pat. No. 4,979,824 describes a device for this purpose. US Pat. No. 4,793,705 describes a detection system for identifying individual molecules in a flow train of the particles in a flow cell. It further describes methods of arranging a plurality of lasers, filters and detectors for detecting different fluorescent nucleic acid base-specific labels. US Pat. No. 4,962,037 also describes a method for detecting an ordered train of labeled nucleotides for obtaining DNA and RNA sequences using an exonuclease to cleave the bases. Single molecule detection on solid supports is also described in Ishikawa, *et al.* (1994) Single-molecule detection by laser-induced fluorescence technique with a position-sensitive photon-counting apparatus, *Jan. J. Appl. Phys.* 33:1571-1576. Ishikawa describes a typical apparatus involving a photon-counting camera system attached to a fluorescence microscope. Lee *et al.* (*Anal. Chem.*, 66:4142-4149, 1994) describes an apparatus for detecting single molecules in a quartz capillary tube. The selection of lasers is dependent on the label and the quality of light required. Diode, helium neon, argon ion, argon-krypton mixed ion, and double Nd:YAG lasers are useful in this invention.

C. Excitation and scanning

In some applications, fluorescent excitation is exerted with a Q-switched frequency doubled Nd YAG laser, which has a KHz repetition rate, allowing many samples to be taken per second. For example, a wavelength of 532 nm is ideal for the excitation of rhodamine. It is a standard device that has been used in the single molecule detection scheme (Smith et al., *Science* 253:1122, 1992). A pulsed laser allows time resolved experiments, which are useful for rejecting extraneous noise. In some methods, excitation can be performed with a mercury lamp and signals from the incorporated nucleotides can be detected with an CCD camera (see, e.g., Unger et al., *Biotechniques* 27:1008, 1999).

Incorporated signals can be detected by scanning the substrates. The substrates can be scanned simultaneously or serially, depending on the scanning method used. The signals can be scanned using a CCD camera (TE/CCD512SF, Princeton Instruments, Trenton, N.J.) with suitable optics (Ploem, J. S., in *Fluorescent and Luminescent Probes for Biological Activity*, Mason, T. W., Ed., Academic Press, London, pp. 1-11, 1993), such as described in Yershov et al. (*Proc. Natl. Acad. Sci.* 93:4913, 1996), or can be imaged by TV monitoring (Khrapko et al., *DNA Sequencing* 1:375, 1991). The scanning system should be able to reproducibly scan the substrates. Where appropriate, e.g., for a two dimensional substrate where the substrates are localized to positions thereon, the scanning system should positionally define the substrates attached thereon to a reproducible coordinate system. It is important that the positional identification of substrates be repeatable in successive scan steps.

Various scanning systems can be employed in the methods and apparatus of the present invention. For example, electro-optical scanning devices described in, e.g., U.S. Pat. No. 5,143,854, are suitable for use with the present invention. The system could exhibit many of the features of photographic scanners, digitizers or even compact disk reading devices. For example, a model no. PM500-A1 x-y translation table manufactured by Newport Corporation can be attached to a detector unit. The x-y translation table is connected to and controlled by an appropriately programmed digital computer such as an IBM PC/AT or AT compatible computer. The detection system can be a model no. R943-02 photomultiplier tube manufactured by Hamamatsu, attached to a preamplifier, e.g., a model no. SR440 manufactured by Stanford Research Systems, and to a photon counter, e.g., an SR430 manufactured by Stanford Research System, or a multichannel detection device. Although a digital signal can usually be preferred, there can be circumstances where analog signals would be advantageous.

The stability and reproducibility of the positional localization in scanning determine, to a large extent, the resolution for separating closely positioned polynucleotide clusters on a two dimensional substrate. Since the successive monitoring at a given position depends upon the ability to map the results of a reaction cycle to its effect on a positionally mapped polynucleotides, high resolution scanning is preferred. As the resolution increases, the upper limit to the number of possible polynucleotides which can be sequenced on a single matrix also increases. Crude scanning systems can resolve only on the order of 1000  $\mu\text{m}$ , refined scanning systems can resolve on the order of 100  $\mu\text{m}$ , more refined systems can resolve on the order of about 10  $\mu\text{m}$ , and with optical magnification systems a resolution on the order of 1.0  $\mu\text{m}$  is available. The limitations on the resolution can be diffraction limited and advantages can arise from using shorter wavelength radiation for fluorescent scanning steps. However, with increased resolution, the time required to fully scan a matrix can increased and a compromise between speed and resolution can be selected. Parallel detection devices which provide high resolution with shorter scan times are applicable where multiple detectors are moved in parallel.

In some applications, resolution often is not so important and sensitivity is emphasized. However, the reliability of a signal can be pre-selected by counting photons and continuing to count for a longer period at positions where intensity of signal is lower. Although this decreases scan speed, it can increase reliability of the signal determination. Various signal detection and processing algorithms can be incorporated into the detection system. In some methods, the distribution of signal intensities of pixels across the region of signal are evaluated to determine whether the distribution of intensities corresponds to a time positive signal.

#### 25 D. Detection of incorporation of multiple fluorescent labels: FRET

In some aspects of the present application, incorporation of different types of nucleotides into a primer is detected using different fluorescent labels on the different types of nucleotides. When two different labels are incorporated into the primer in close vicinity, signals due to fluorescence resonance energy transfer (FRET) can be detected. FRET is a phenomenon that has been well documented in the literature, e.g., in T. Foster, Modern Quantum Chemistry, Istanbul Lectures, Part III, 93-137, 1965, Academic Press, New York; and Selvin, "Fluorescence Resonance Energy Transfer," Methods in Enzymology 246: 300-335, 1995. In FRET, one of the fluorophores (donor) has an emission spectrum that overlaps the excitation spectrum of the other fluorophore (acceptor) and transfer of energy takes place

from the donor to the acceptor through fluorescence resonance energy transfer. The energy transfer is mediated by dipole-dipole interaction. Spectroscopically, when the donor is excited, its specific emission intensity decreases while the acceptor's specific emission intensity increases, resulting in fluorescence enhancement.

5 Detection of single molecule FRET signal reveals sequence information and facilitates interpretation of the sequencing data. Detection of FRET signal in the present invention can be performed accordingly to various methods described in the art (e.g., US Patent No. 5,776,782). FRET has been used to studying various biological activities of biomacromolecules including polynucleotides. For example, Cooper et al. disclosed  
10 fluorescence energy transfer in duplex and branched DNA molecules (Biochemistry 29: 9261-9268, 1990). Lazowski et al. reported highly sensitive detection of hybridization of oligonucleotides to specific sequences of nucleic acids by FRET (Antisense Nucleic Acid Drug Dev. 10: 97-103, 2000). Methods for nucleic acid analysis using FRET were also described in US Patent Nos. 6,177,249 and 5,945,283. Efficacy of using FRET to detect  
15 multiple nucleotides incorporation into single polynucleotide molecules is also exemplified in Example 8 of the present application.

Any of a number of fluorophore combinations can be selected for labeling the nucleotides in the present invention for detection of FRET signals (see for example, Pesce et al., eds, Fluorescence Spectroscopy, Marcel Dekker, New York, 1971; White et al.,  
20 Fluorescence Analysis: A practical Approach, Marcel Dekker, New York, 1970; Handbook of Fluorescent Probes and Research Chemicals, 6th Ed, Molecular Probes, Inc., Eugene, Oreg., 1996; which are incorporated by reference). In general, a preferred donor fluorophore is selected that has a substantial spectrum of the acceptor fluorophore. Furthermore, it may also be desirable in certain applications that the donor have an excitation maximum near a  
25 laser frequency such as Helium-Cadmium 442 nm or Argon 488 nm. In such applications the use of intense laser light can serve as an effective means to excite the donor fluorophore. The acceptor fluorophore has a substantial overlap of its excitation spectrum with the emission spectrum of the donor fluorophore. In addition, the wavelength of the maximum of the emission spectrum of the acceptor moiety is preferably at least 10 nm greater than the  
30 wavelength of the maximum of the excitation spectrum of the donor moiety. The emission spectrum of the acceptor fluorophore is shifted compared to the donor spectrum.

Suitable donors and acceptors operating on the principle of fluorescence energy transfer (FET) include, but are not limited to, 4-acetamido-4'-isothiocyanatostilbene-2,2'-disulfonic acid; acridine and derivatives: acridine, acridine isothiocyanate; 5-(2'-



aminoethyl)aminonaphthalene-1-sulfonic acid (EDANS); 4-amino-N-[3-vinylsulfonyl]phenyl]naphthalimide-3,5 disulfonate; N-(4-anilino-1-naphthyl)maleimide; anthranilamide; BODIPY; Brilliant Yellow; coumarin and derivatives: coumarin, 7-amino-4-methylcoumarin (AMC, Coumarin 120), 7-amino-4-trifluoromethylcoumarin (Coumaran 151); cyanine dyes; cyanosine; 4',6-diaminidino-2-phenylindole (DAPI); 5', 5"-dibromopyrogallol-sulfonaphthalein (Bromopyrogallol Red); 7-diethylamino-3-(4'-isothiocyanatophenyl)-4-methylcoumarin; diethylenetriamine pentaacetate; 4,4'-diisothiocyanatodihydro-stilbene-2,2'-disulfonic acid; 4,4'-diisothiocyanatostilbene-2,2'-disulfonic acid; 5-[dimethylamino]naphthalene-1-sulfonyl chloride (DNS, dansylchloride); 4-dimethylaminophenylazophenyl-4'-isothiocyanate (DABITC); eosin and derivatives: eosin, eosin isothiocyanate, erythrosin and derivatives: erythrosin B, erythrosin, isothiocyanate; ethidium; fluorescein and derivatives: 5-carboxyfluorescein (FAM), 5-(4,6-dichlorotriazin-2-yl)aminofluorescein (DTAF), 2',7'-dimethoxy-4'5'-dichloro-6-carboxyfluorescein (JOE), fluorescein, fluorescein isothiocyanate, QFITC, (XRITC); fluorescamine; IR144; IR1446; Malachite Green isothiocyanate; 4-methylumbelliferoneortho cresolphthalein; nitrotyrosine; pararosaniline; Phenol Red; B-phycoerythrin; o-phthalaldehyde; pyrene and derivatives: pyrene, pyrene butyrate, succinimidyl 1-pyrene; butyrate quantum dots; Reactive Red 4 (Cibacron™ Brilliant Red 3B-A) rhodamine and derivatives: 6-carboxy-X-rhodamine (ROX), 6-carboxyrhodamine (R6G), lissamine rhodamine B sulfonyl chloride rhodamine (Rhod), rhodamine B, rhodamine 123, rhodamine X isothiocyanate, sulforhodamine B, sulforhodamine 101, sulfonyl chloride derivative of sulforhodamine 101 (Texas Red); N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA); tetramethyl rhodamine; tetramethyl rhodamine isothiocyanate (TRITC); riboflavin; rosolic acid; terbium chelate derivatives; Cy 3; Cy 5; Cy 5.5; Cy 7; IRD 700; IRD 800; La Jolla Blue; phthalocyanine; and naphthalocyanine.

\*\*\*

Many modifications and variations of this invention can be made without departing from its spirit and scope. The specific embodiments described below are for illustration only and are not intended to limit the invention in any way.

EXAMPLESExample 1 Basic Materials and Methods

## 5 1. Materials and Reaction Reagents

## (1) Solutions and buffers

RCA: H<sub>2</sub>O:NH<sub>4</sub>OH:H<sub>2</sub>O<sub>2</sub> (6:4:1) boiling for an hour.

PEI: PolyEthylenImine (Sigma P-3143) (positive charged)

PALL: Poly(allylamine hydrochloride) (Sigma 283223)

10 PACr: Poly(acrylic acid, sodium salt) (Sigma 416045) (negative charged)

EDC: 9.6mg/ml; 50mM (x10) 1-[3-(Dimethylamino)propyl]-3-ethylcarbodiimide, hydrochloride), Activator for the BLCPA (Sigma- 161462)

BLCPA: EZ-Link Biotin LC-PEO-Amine (Pierce 21347)

Stock solution 50mM in MES 10mM (21mg/ml) (x10)

15 Streptavidin plus - 1mg/ml in Tris. PROzyme, Code: SA20 (x10)

## Buffers:

MES (N-morpholinoethanesulfonic acid) PH 5.5 1M (100x)

TRIS 10mM

20 TRIS-MgCl<sub>2</sub> 10mM Tris, 100mM MgCl<sub>2</sub> (x1)TKMC (10mM Tris• HCl, 10mM KCl, 10mM MgCl<sub>2</sub>, 5mM Ca Cl<sub>2</sub>, pH 7.0)EcoPol : 10mM Tris• HCl, 5mM MgCl<sub>2</sub> , 7.5 mM DTT pH @ 25<sup>0</sup>C; buffer come with the polymerase at (x10)

## 25 (2) Other materials and reagents

Nucleotides: dTTP, dGTP, dATP, and dCTP-Cy3 at 10μM concentration

Polymerase: a) Klenow Polymerase I (5 units/μl), New England BioLabs Cat. 210S

b) Klenow -exo, New England BioLabs Cat. 212S

30 c) TAQ

d) Sequenase

Hybridization Chamber: Sigma H-1409

Polynucleotide templates and primers:

35 7G: Biotin - 5'-tcagtcatca gtcacagtc atcagtcac agtcacagtc catcagtcac  
cagtcacag tcacagtcac tcagtcacac gtcacACAC GGAGGTTCTA - 3' (SEQ ID NO:1)

Primer p7G: 5'- TAGAACCTCCGTGT - 3' (SEQ ID NO:2); the primer can be labeled with Cy5 or Cy3.

40 ID NO:3) Mu50: Biotin 5'- ctccagcgtgtttatctctgcgagcacaatgcctgcgcatccgccagc 3' (SEQ

Cy5 labeled primer (PMu50Cy5): Cy5 5' - gctggcggatgac - 3' (SEQ ID NO:4)

7G7A - Biotin-5'-

45 tttGcttcttAttctttGcttcttAttctttGcttcttAttctttGcttcttAttctttGcttcttAttctttGcttcttAttcttACACGGA  
GGTTCTA - 3' (SEQ ID NO:5)

6TA6CG: Biotin- 5' –

ccAtttttGccccccAtttttGccccccAtttttGcc...ccAtttttGccccccAtttttA-CACGGAGGTTCTA -  
3', (SEQ ID NO:6)

## 5 2. Substrate treatment and template attachment

A fused silica microscope slide (1 mm thick, 25x75 mm size, Esco Cat. R130110) was used to attach DNA templates. The slides was first cleaned with the RCA method as described above and in WO 01/32930. Multilayer of polyallylamine /polyAcrylic were absorbed to the slide. An EZ link connector was then attached to the slides as follows:  
10 the slide was dried, scratched with diamond pencil, and then covered with a hybridization chamber. 120 µl of a mixture of 1:1:8 EDC: BLCPA: MES (50mM EDC, 50mM BLCPA, 10mM MES) was applied to each slide. Following incubation for 20 minutes, 120 µl of Streptavidin Plus diluted to 0.1mg/ml was added to the slide. After 20 min of incubation, the slide was washed with 200µl of Tris 10mM.

15 Preparation of 10pM Oligo: the 7G oligonucleotide template (SEQ ID NO:1) was pre-hybridized with Cy5-labeled primer (SEQ ID NO:2) (in stock at 7µM) in TRIS-MgCl<sub>2</sub> buffer. The treated slide was examined for contamination with the TIR microscope. 200µl of the oligonucleotide/primer mixture was applied to each slide. Following incubation for 10 min, the slide was washed with 200µl ml of Tris 10mM.

20 Addition of nucleotides and polymerase: nucleotides dTTP, dATP, dGTP, and Cy3-dCTP each of 20-100nM were mixed in the ECOPOL buffer. 1 µl Klenow 210S from stock solution (kept in -20°C) was added to 200 microliters of the nucleotide mixture. 120µl of the mixture was then added on each slide. After incubation for 0 to 30 min (for different experiments), the slide was examined with the TIR microscope. Unless otherwise noted, all  
25 reactions were performed at room temperature, while the reaction reagents were kept at 4°C or -20°C. The primer/oligonucleotide hybridization reaction was carried out with a thermocycler machine.

Single molecule resolution was achieve by using very low concentration of the polynucleotide template which ensured that only one template molecule is attached to a  
30 distinct spot on the slide. Single molecule attachment to a distinct is also confirmed by the observation of single bleaching pattern of the attached fluorophores. In the reaction described above, a concentration of about 10pM of a 80-mer oligonucleotide template was

used for immobilizing to the slide. The space between different DNA molecules attached to the surface slide was measured at a few micrometers.

### 3. Imagine with single molecule resolution

As illustrate in Figure 1, the single stranded oligonucleotide template (SEQ ID NO:1) primed with a Cy5 labeled primer sequence (SEQ ID NO:2) was immobilized at a single molecule resolution to the surface of a silica slide using a biotin-streptavidin bond. The surface is coated with polymers on which biotin (EZ link) is tethered. The oligonucleotide template, with a biotin molecule attached to one of its ends, was able to attach to the streptavidin-linked surface. The slide surface was negatively charged which helps to repeal unbound nucleotides. The DNA is specifically attached to the surface by its 5' side, meaning that the primer -which the polymerase extends- is away from the surface.

The template and incorporation of labeled nucleotides were visualized by fluorescence imaging. Location of the oligonucleotide was monitored by fluorescence from the Cy5 labeled primer (SEQ ID NO:2). Incorporation of nucleotides was detected because the nucleotides were labeled with Cy3. After incorporation, the incorporated labels were illuminated. Illumination of Cy3 was at a wavelength of 532nm. Following a typical time of a few seconds of continued illumination, the signals were bleached, typically in a single step.

As shown in Figure 2, imaging of fluorescent signals with single molecule resolution was enabled with surface illumination by total internal reflection (TIR). Ishijima et al. (Cell 92:161-71, 1998) showed that it is possible to observe the fluorescence of single molecules immobilized to a surface in a wet environment even when there are free molecules in the solution. Here, the TIR was facilitated by a dove prism coupling of the laser beam to the silica slide surface. An upright microscope with an immersion oil objective was used to image the surface with an intensified CCD (PentaMax). A filter set (Chroma) was used to reject the illumination frequency and let the fluorescence frequency to reach the ICCD.

#### Example 2 Test for Specific Attachment of Template Molecules to Substrate Surface

This experiment was performed to determine whether the polynucleotide templates are attached to the surface as desired. Figure 3 shows that streptavidin is required for binding the template to the surface and hence detection of incorporated fluorescence signal. The left panel shows that there is no fluorescence signal when only streptavidin-attached surface but no fluorescent labels were present. The middle panel shows that there is no incorporated fluorescent signals when no streptavidin was present on the surface to attach

biotin-labeled oligonucleotide template, even though Cy5-labeled primer was present. The right panel shows that detection of incorporated fluorescent signal when the streptavidin-attached surface, labeled primers, and biotin-labeled oligonucleotide template were present.

5    Example 3.    Determining Processivity of DNA Polymerase in the Presence of Labeled Nucleotides

To determine whether the DNA polymerase accurately incorporates labeled nucleotides into the template, a bulk extension experiment was performed in a test tube rather than on the surface of a substrate. As shown in Figure 5, the results indicate that the  
10    polymerase incorporate all the labeled nucleotides into the correct positions. In this experiment, incorporation of dCTP-Cy3 and a polymerization terminator, ddCTP, were detected using a 7G DNA template (a DNA strand having a G residue every 7 bases; SEQ ID NO:1). The annealed primer was extended in the presence of non-labeled dATP, dGTP, dTTP, Cy3-labeled dCTP, and ddCTP. The ratio of Cy3-dCTP and ddCTP was 3:1. The  
15    reaction products were separated on a gel, fluorescence excited, and the signals detected, using an automatic sequencer ABI-377. The results reveal that incorporation of Cy3-dCTP did not interfere with further extension of the primer along the 7G oligomer template.

Figure 5 shows fluorescence intensity from primer extension products of various lengths which were terminated by incorporation of ddCTP at the different G residues  
20    in the 7G oligomer template (SEQ ID NO:1). The first band is the end of the gel and should not be counted as it is in the very beginning of the gel. The full length of the template is 100 residues. The first band (marked "1" in the graph) corresponds to extension products which were terminated by incorporation of non-labeled ddCTP at the second G residue (position 27) and has incorporated Cy3-dCTP at the first G residue (position 20). Similarly, the tenth band  
25    (marked "10" in the graph) represents extension products which were terminated by incorporation of non-labeled ddCTP at the 10th G residue (position 90) and has incorporated Cy3-dCTP at the previous G residue (i.e., positions 20, 27, 34, 41, 48, 55, 62, 69, 76, and 83). The results showed a nice agreement between the expected positions for Cy3 incorporation in the polynucleotide template and the positions of the fluorescence intensity bands.

30

Example 4.    Detection of single nucleotide incorporation by TIR

Total internal reflection (TIR) fluorescence microscopy allows detection of real-time incorporation of labeled nucleotide into single immobilized polynucleotide template. This illumination method reduce the background from the sample by illuminating

only a thin layer (e.g., in the order of 150nm) near the surface. Even in the presence of free dyes in the solution (up to 50nM), single molecules can be observed. Using TIR, we visualized single molecules of labeled nucleotide bound to DNA in the presence of up to 50nM free dye in solution. Though this concentration is low compared to the concentration  
5 needed for a high rate of incorporation of nucleotides by the DNA polymerase, it was sufficient for its operation.

### 1. Optical setup

The lasers source is shown in Figure 2, the light sources (e.g., laser) are  
10 coupled to the surface by prism. The surface is imaged by a regular 1.3NA microscope objective onto an Intensified CCD (Pentamax). A fluorescent filter in the optical way block the laser intensity and allow the fluorescent signals from the dye molecules pass through(Chroma filters). Optionally, the camera and the shutters for the lasers are controlled by the computer.

### 2. Illumination

As shown in Figure 6, TIR illumination of polynucleotide-attached slide produced a low background and allowed detection of signals only from immobilized labels. The refraction index of the fused silica glass and the oil beneath the surface is about 1.46.  
20 The refraction index of the liquid above the glass is about 1.33 to 1.35. At the interface of the glass and the water the illumination ray was refracted. If the illumination is very shallow, 70-75 degree from the surface orthogonal, the refracted light was reflected back and not continued in the liquid phase as the critical angel for total internal reflection is about 65-67 degrees ( $\text{TetaCritical} = \sin^{-1}(n1/n2)$ ).

25 The illumination process, called evanescent illumination, leaves a decay field near the interface which illuminates only about 150 nm into the liquid phase. Fluorophores dyes can be excited by this field. So only the dyes which are near the surface will emit. Furthermore, free labeled nucleotide molecules in the solution will move around due to Brownian motion. The fast movement of these free molecules produces only a smear signal  
30 because the integration time is in the order of hundred millisecond. Thus, the total internal reflection illumination leads to a low back ground from the free molecules, and only signals from the immobilized dyes are detected.

### 3. Detection of single molecules

Figures 6 shows detection of signals from single Cy3 molecule with no free dye in solution versus signals from single Cy3 molecule with background of 15nM Cy3 in solution. Fluorescence image from incorporation of Cy3 labeled nucleotide is shown in the upper panels. The signals tend to bleach in a single step, see the upper graph. When there are free labeled nucleotides in the solution (15nM free dye), the background signal is stronger (lower right panel) than the background signal in the absence of free labeled nucleotides in the solution. But the signal from the incorporated single molecule can still be detected. The ability to detect single molecule in the presence of free dye enables one to follow incorporation of nucleotide into an immobilized DNA template in real time.

The upper left panel of Figure 6 showed typical images of single molecules (see the bright spots). When the intensity of a spot is traced in real time (upper right panel), one can see that it appears (incorporation event or sticking to the surface event) and disappears (bleaching or detaching event). The same results are also illustrated in the middle long thin panel of Figure 6. This panel shows successive images of a small area around the spot that was being traced. The fluorescent signal appeared and disappeared after every few seconds (every frame is a second exposure).

**Example 5. Determining Nucleotide Incorporation Based on Correlation of Fluorescence Spots**

A correlation was observed between the position of the immobilized DNA template on the surface (indicated by the fluorescently labeled primer) and the incorporation of nucleotide to the surface. In Figure 4, image of the immobilized DNA which was hybridized to the Cy5 labeled primer was shown in the upper two panels (the middle panel is a magnified image of a small area in the left panel). The small dots in the image represent likely positions of the DNA templates immobilized on the surface. The fluorescence signals were then bleached out by a long radiation (about 1 minute) at 635nm with a 10mW laser diode. Subsequently, the polymerase and the nucleotides (including the Cy3-labeled dCTP) were added, and the mixture incubated at room temperature for about an hour. After washing, a second image of the surface was taken. This time a new set of fluorescence-labeled points appeared (see lower left two panels). The results indicate that the two sets of fluorescently-labeled points are correlated (see right panel). It is noted that the significant overlap (about 40%) between DNA primer location (Cy5) and dCTP Incorporation location (Cy3) cannot be a random result. Under the concentrations of labeled DNA primers used in the experiment, the probability for this correlation to occur randomly calculated to be about

10<sup>-50</sup>. Rather, the correlation is due to incorporation of the Cy3 labeled nucleotides into the immobilized, Cy5 labeled primer.

Incorporation of labeled nucleotide into the immobilized template is also demonstrated by the multi-incorporation data shown in Figure 7. When the intensity of the spots in Figure 4 were measured, a multistep bleaching is observed (Figure 7, upper left panel). Simulation of the multiple bleaching is shown in the upper right panel. The results are what should be expected if few molecules are located in the same place up to the optical resolution. This indicates that the polymerase can incorporate a few labeled nucleotides into the same DNA template. In a control experiment, ddATP, dCTP-Cy3 and dGTP were used to extend Cy5-labeled primer PMu50Cy5, Cy5 5' - gctggcggatgac - 3' (SEQ ID NO:4) along the Mu50 oligonucleotide template (SEQ ID NO 3). This allows only one Cy3-labeled nucleotide to be incorporated into the primer because the first codon in the template sequence after the primer is CGT. Incorporation of ddATP immediately after the incorporation of dCTP-Cy3 terminates the elongation. As shown in the lower right panel, there is no multibleaching.

It is noted that because the concentration of the DNA template on the surface was so low, it is unlikely that more than one copy of the DNA template were present on each spot. Further, multiple bleaching is not common when the polymerase was not present (data not shown). In particular, there is no correlation between primer location and fluorescence signal from the surface when the polymerase was not present (see, e.g., Figure 13, middle panel).

#### Example 6. Dynamics of Nucleotide Incorporation

Figure 8 shows a time course of incorporation events during the DNA polymerase reaction. In this experiment, the DNA template and Cy5-labeled primer complex was immobilized to the substrate surface as described above, and its position was imaged. The DNA Polymerase was then added along with the nucleotides of which one was labeled with Cy3.

As indicated in the figure, the substrate was imaged every 10 sec, with a 1 sec exposure. Every spot with immobilized DNA template (as indicated by the labeled primer) was monitored as a function of time. A series of small images of these spots were placed along a strip resulting in a movie showing the "activities" at each point.

Repeated incorporation of nucleotide into the DNA template was shown in Figure 9. Using more dyes will enable us to read the sequence of the DNA directly in an



asynchronous manner. Figure 9 shows the dynamic incorporation events at 8 different spots. The digital information recorded in these movies indicate that repeated incorporation events occurred at various time points. The data also demonstrated the feasibility of monitoring primer extension activities on single DNA molecules.

Figure 10 shows a histogram of the number of incorporation events on single spots and a histogram of the time between incorporation events. From the histograms one can see that a few nucleotides were incorporated into single DNA molecules. The low numbers of events in which more than three nucleotides were incorporated indicate that there is some mechanism that prevents high number of incorporation into the DNA under the experimental conditions. The reason could be that photo-damage to the DNA in the surrounding area of the illuminated dye might produce toxic radicals. Changing the reaction conditions and reagents could increase the numbers of incorporated nucleotides dramatically.

#### Example 7 Base-by-base Sequence Analysis

This experiment was performed to confirm selectivity of the polymerase and to illustrate feasibility of determining the sequence of a polynucleotide template with base-by-base scheme.

First, fidelity of the polymerase in incorporation was confirmed by analyzing correlation between location of immobilized primer and location of nucleotide incorporation with a correlation graph. Figure 11 shows correlation between primer location and polymerase activity location. The position of each point was determined with a sub pixel resolution. Images for the primer location and the incorporation position were taken first. If there is a correlation between the two, there is a pick in the correlation graph. Otherwise no pick was observed. As shown in the figure, the two images correlate with each other.

Results demonstrating base-by-base analysis of the sequence of a immobilized template at single molecule resolution is shown in Figure 12. The data indicated that at least two bases of the template were determined by flowing in and out reagents along with different types of labeled nucleotides (e.g., dCTP-Cy3, dUTP-Cy3, etc.). Here, a 6TA6GC oligonucleotide template (SEQ ID NO:6) was immobilized to the fused silica slide. A Cy3-labeled p7G primer (SEQ ID NO:2) was annealed to the template. As illustrated in the Figure, the primer was first extended up to the A residue with non-labeled dATP nucleotides. Then, dUTP-Cy3 nucleotide was incorporated and imaged. Images taken at this time show high correlation (see the upper left correlation graph). After bleaching the dyes, dCTP-Cy3 was applied to the sample. Images taken at this time show low correlation (see the lower left

correlation graph). Thereafter, non-labeled dGTP was added to fill the CCCCC gap till the G residue in the sequence. At this time, incorporation of a dCTP-Cy3 nucleotide was examined again. This time there was a correlation between the dCTP-cy3 positions and the primer positions in general, and in particular there was a correlation with the position of the incorporated dUTP in the first incorporation cycle. Thereafter, dUTP-Cy3 was added. Correlation was found between the labeled primer position and signal from dUPT-Cy3, but no correlation was found between the new dUPT-Cy3 positions and the position that has incorporated dUTP in the first incorporation cycle (lower right graph). The interpretation is that not all the primers were extended in the first dUTP incorporation cycle, that those which did not get extended could incorporate dUTP in the second incorporation cycle, and that those which did incorporate dUTP in the first cycle could not incorporate dUTP again in the second cycle. The results indicate that on those spots which have incorporated the first U residue there were also incorporations of a C but not a U residue. Thus, identity of a second base can be determined with the experimental scheme, although the yield for the second base (upper right graph) was not as good as for the first base (upper left graph).

In a control experiment, after filling in with A residues, dCTP-Cy3 (wrong nucleotide for the first base) was added. Correlation between Cy3-labeled primer position and C-Cy3 was low (data not shown). In another control, after filling in the string of A residues, the U residue, G residues, and U-Cy3 (wrong residue for the second base) was added. The correlation observed from the results in this experiment was low (at the noise level; data not shown). Using different oligonucleotide templates, the experiment scheme was repeated for successive incorporations of other combinations of two or more nucleotides (data not shown). The results confirmed correct incorporation of the first labeled nucleotide with high signal-to-noise ratio and subsequent incorporations of more nucleotides with a relatively lower signal-to-noise ratio. Taken together, these data indicate that the observed results (e.g., as shown in Figure 12) are not due to artifacts, but rather demonstrate efficacy of base-by-base analysis of the experimental scheme.

#### Example 8. Two Color Incorporation: Fluorescence Resonance Energy Transfer

This experiment demonstrate incorporation of two different fluorescent labels into the same immobilized polynucleotide template through detection of fluorescence resonance energy transfer (FRET). In this experiment, two fluorescent labels were used (Cy5 and Cy3), and FRET from dUTP-Cy3 (donor) to dCTP-Cy5 (acceptor) was examined at the single molecule level as shown in Figure 13.

Image of the DNA template with the labeled primer is shown in the left panel. Detection of FRET after incorporation of the two labels is provided in the right image. Correlation between the template location and the incorporation signals is shown in the middle graph. As indicated, there is a high correlation between the template location and the incorporated nucleotide location. A control experiment was performed in which no polymerase is present. Results from the control experiment produced a low correlation between the template location and location of labeled nucleotides. FRET experiment provides particularly high signal to noise ratio as there is almost no signal from nonspecific incorporation of dyes to the surface.

When the two labels were incorporated into a primer at close vicinity, i.e., at a few nanometers apart, a single molecule FRET signal was detected (Figure 14). To detect the FRET signal, the optic setup was altered. A image splitter was added so that the same area was imaged twice (Optical Insights LTD, micro imager device). In one channel, a fluorescence filter detected only the donor (cy3) fluorescence. In the other channel, a filter for the acceptor (Cy5) was placed. With this setup individual spots were examined after incorporation. Figure 15 further indicates that the FRET detection scheme allows measurement of incorporation rate with a nice signal to noise ratio.

WHAT IS CLAIMED IS:

1. A method of analyzing sequence of a target polynucleotide, comprising:
  - (a) providing a primed target polynucleotide immobilized to a surface of a substrate; wherein the target polynucleotide is attached to the surface with single molecule resolution;
  - (b) adding a first fluorescently labeled nucleotide to the surface of the substrate under conditions whereby the first nucleotide attaches to the primer, if a complementary nucleotide is present to serve as template in the target polynucleotide;
  - (c) determining presence or absence of a fluorescence signal on the surface where the target polynucleotide is immobilized, the presence of a signal indicating that the first nucleotide was incorporated into the primer, and hence the identity of the complementary base that served as a template in the target polynucleotide; and
  - (d) repeating steps (b)-(c) with a further fluorescently labeled nucleotide, the same or different from the first nucleotide, whereby the further nucleotide attaches to the primer or a nucleotide previously incorporated into the primer.
2. The method of claim 1, wherein step (a) comprises providing a plurality of different primed target polynucleotides immobilized to different portions of the substrate.
3. The method of claim 1, wherein steps (b)-(c) are performed at least four times with four different types of labeled nucleotides.
4. The method of claim 1, wherein steps (b)-(c) are performed until the identity of each base in the target polynucleotide has been identified.
5. The method of claim 1, further comprising an additional step of removing the signal after step (c).
6. The method of claim 1, wherein the presence or absence of a fluorescence signal is determined with total internal reflection fluorescence (TIRF) microscopy.

7. The method of claim 1, wherein the target polynucleotide is primed with a fluorescently labeled primer.
8. The method of claim 1, wherein the first and further nucleotide are labeled with the same fluorescent label.
9. The method of claim 1, wherein said the substrate is a fused silica slide.
10. The method of claim 9, wherein said surface is coated with a polyelectrolyte multilayer (PEM).
11. The method of claim 10, wherein said PEM is terminated with a polyanion.
12. The method of claim 11, wherein said polyanion bears pendant carboxylic acid groups.
13. The method of claim 12, wherein said target polynucleotide is biotinylated, and said surface is coated with streptavidin.
14. The method of claim 13, wherein said surface is coated with biotin prior to coating with streptavidin.
15. The method of claim 14, wherein said surface is coated with a polyelectrolyte multilayer (PEM) terminated with carboxylic acid groups prior to attachment of biotin.
16. The method of claim 1, wherein said removing or reducing is by photobleaching.
17. The method of claim 1, wherein the substrate is in fluid communication with a microfluidic device, wherein the first and further labeled nucleotides are added to or removed from the substrate through the microfluidic device.
18. The method of claim 17, wherein the microfluidic device comprises
  - (a) a flow cell comprising the substrate; and

(b) an inlet port and an outlet port, said inlet port and outlet port being in fluid communication with said flow cell for flowing fluids into and through said flow cell.

19. The method of claim 18, wherein the substrate is a microfabricated synthesis channel.

20. The method of claim 17, furthering comprising a light source to illuminate the surface of said substrate and a detection system to detect a signal from said surface.

21. The method of claim 17, further comprising an appropriately programmed computer for recording identity of a nucleotide when said nucleotide becomes incorporated into the target polynucleotide.

22. A method of analyzing sequence of a target polynucleotide, comprising:

(a) providing a primed target polynucleotide immobilized to a surface of a substrate; wherein the target polynucleotide is attached to the surface with single molecule resolution;

(b) adding four types of nucleotides to the surface of the substrate under conditions whereby nucleotides attach to the primer dynamically, when complementary nucleotides are present in the target polynucleotide; and

(c) monitoring in a time course of incorporation of fluorescent signals into the immobilized primer.

23. The method of claim 22, wherein monitoring of fluorescent signal incorporation into the immobilized primer is by taking images in a time course with monitored with total internal reflection fluorescence microscopy.

24. The method of claim 23, wherein the images are taken at a rate faster than the rate at which nucleotides are incorporated into the primer.

25. The method of claim 23, wherein nucleotide concentrations are low at each time point when an image is taken.

26. The method of claim 25, wherein nucleotide concentrations are alternated by fluid exchange with a microfluidic device.

27. The method of claim 22, wherein all four types of nucleotides are each labeled with a different label.

28. An apparatus for analyzing the sequence of a target polynucleotide, comprising:

(a) a flow cell comprising a substrate for immobilizing the target polynucleotide with single molecule resolution;

(b) an inlet port and an outlet port, said inlet port and outlet port being in fluid communication with said flow cell for flowing fluids into and through said flow cell;

(c) a light source for illuminating the surface of the substrate; and

(d) a detection system for detecting a signal from said surface.

# Molecular Setup

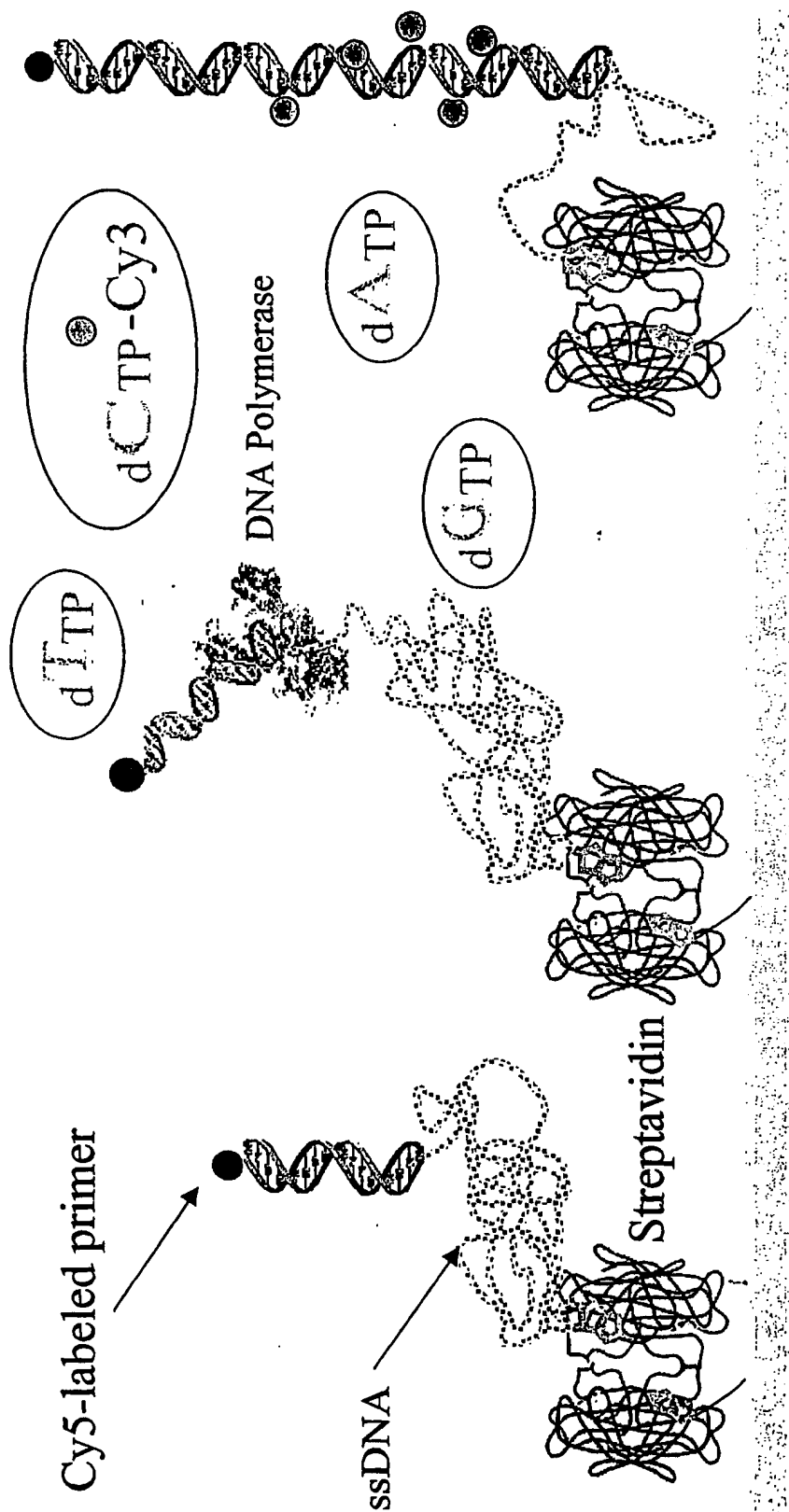


FIG. 1



# TIR Optical Setup

## Total Internal Reflection

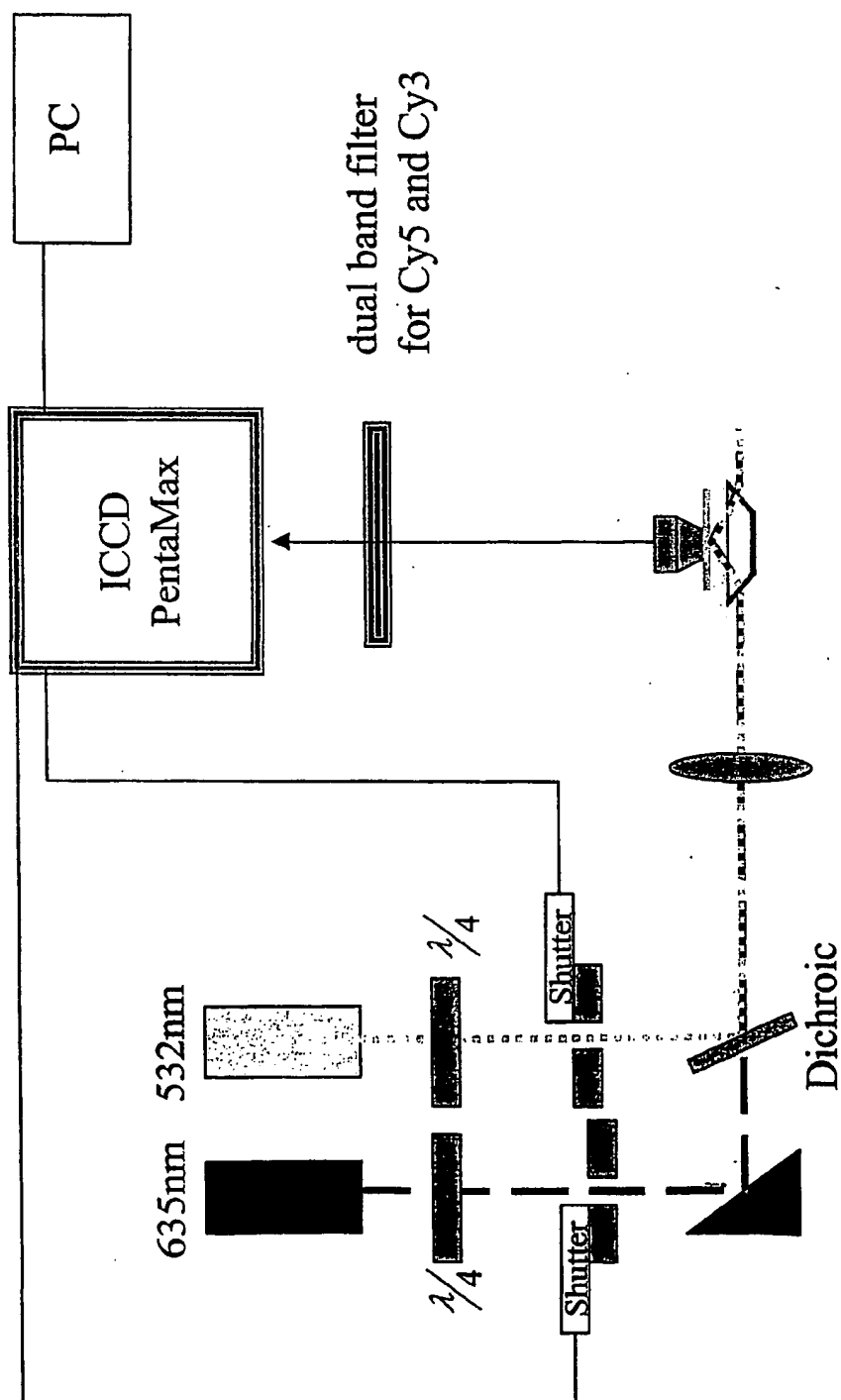
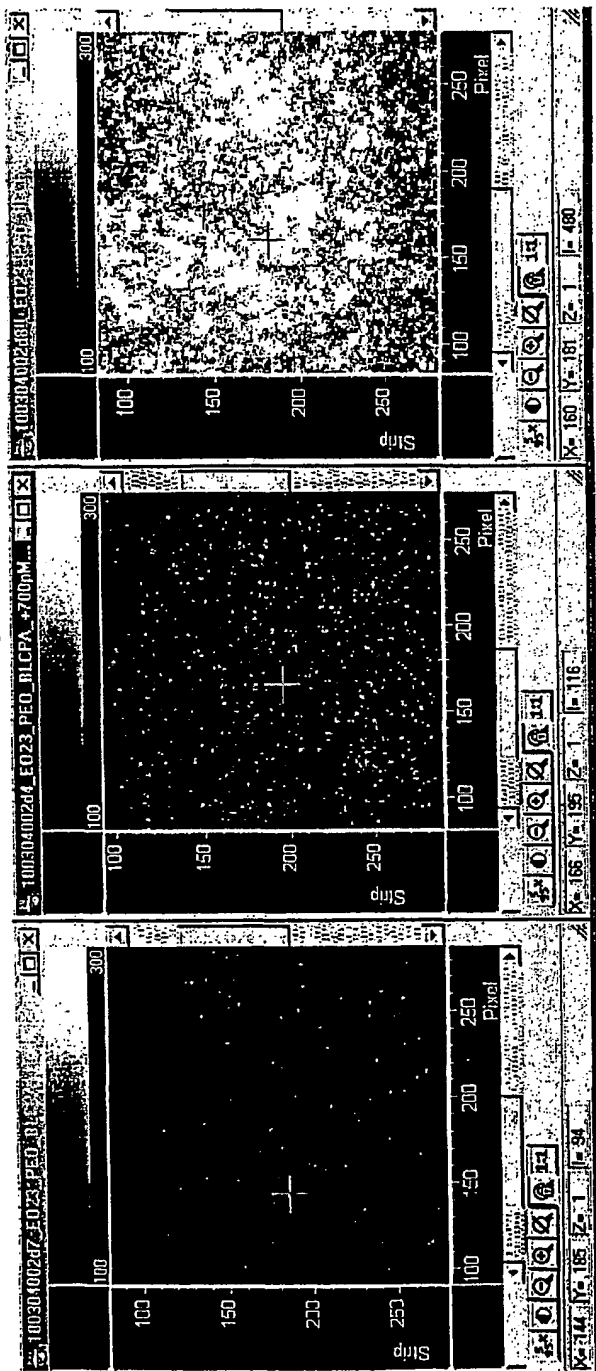


FIG. 2

# Streptavidin Is Required to Bind the DNA Template.



Streptavidin  
Without dye

0.7nM Cy5-primer+  
Biotin\_Oligo,  
Without Streptavidin  
on the surface

0.7nM Cy5-primer+  
Biotin\_Oligo,  
With Streptavidin  
on the surface

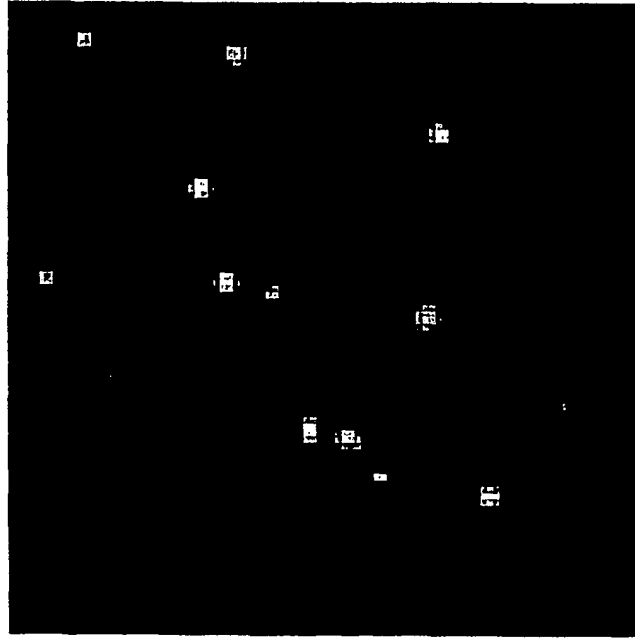
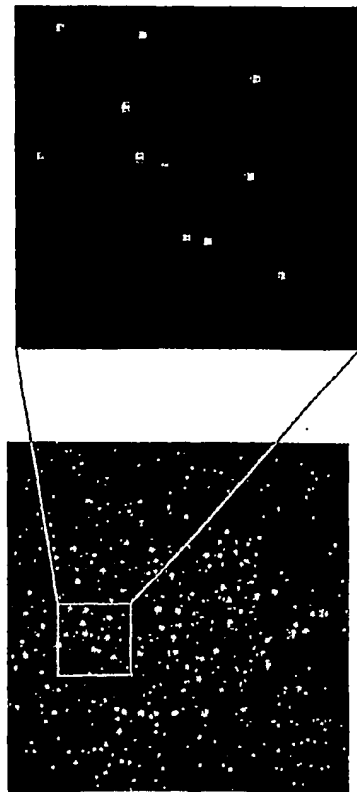
FIG. 3

# The Polymerase Is Active on the Surface

DNA primer location (Cy5)



dCTP Incorporation location (Cy3)



About 40% of the red spots map on to green spots. With these concentrations the probability for this correlation to occur randomly calculated to be about 10<sup>-50</sup>

Fig. 4

# Checking the Incorporation of Labeled Nucleotide in the Bulk

**Biotin-5'-TCAGTCATCAGTCATCAGTCATCAGTCATCAGTCATCAGTCATCAG-**

	12	11	10	9	8	7	6	5	
-TCATCAGTCA	TCA	GTCAGTCA	TCA	TCA	GTCAGTCA	TCA	TCA	CACGAGTTCTA-3'	
4	3	2	1	0				3'-TGTCCTCCAAGAT-5'-	

Cy3-dCTP:ddCTP    3:1

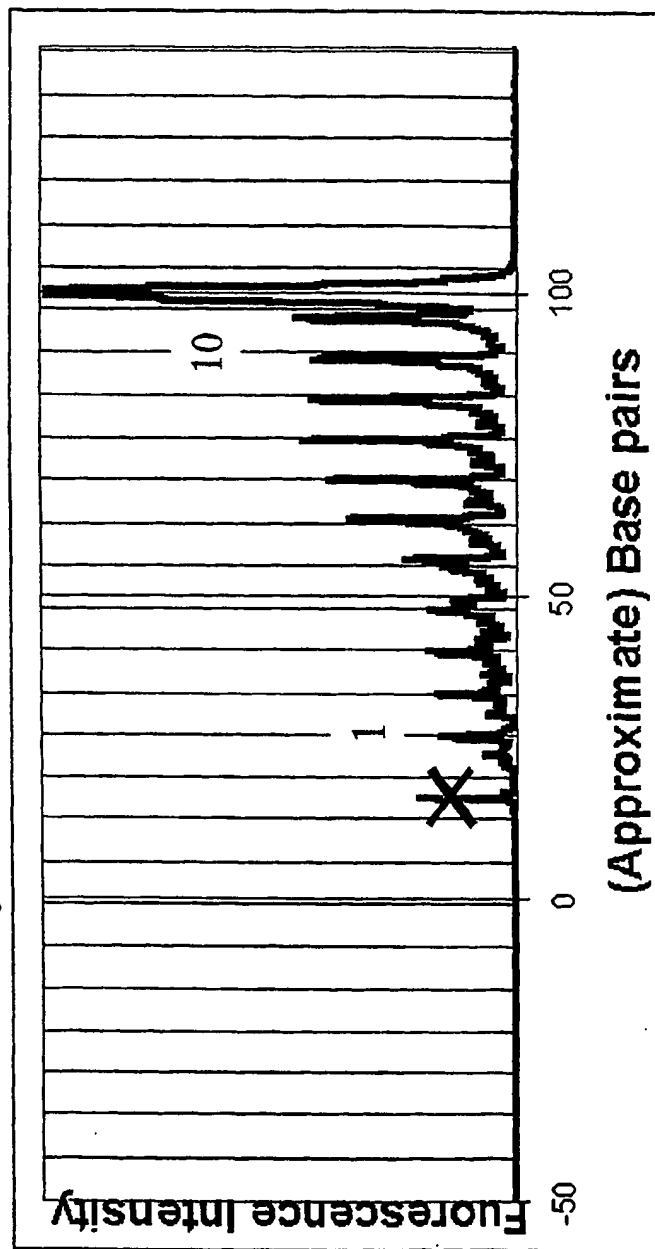


FIG. 5

# Detection of Single Molecules

- clean controls
- bleaching in one step
- number of objects correlates with with concentration of fluorophores
- histogram has only one peak

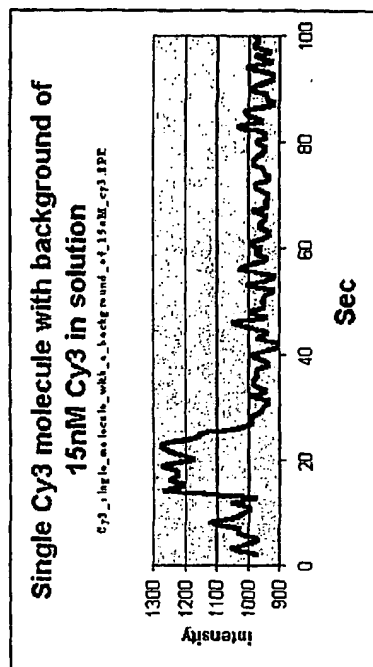
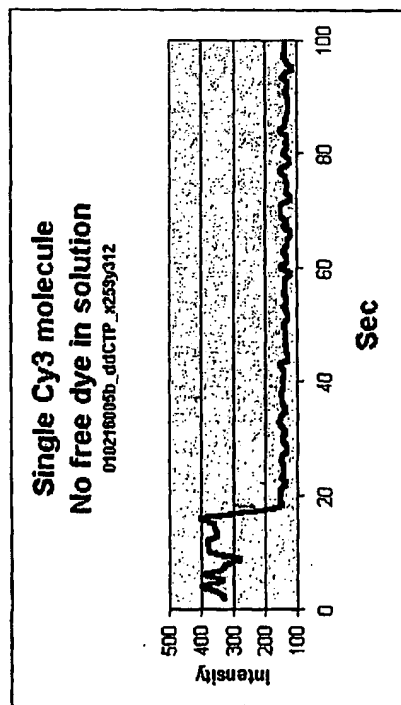
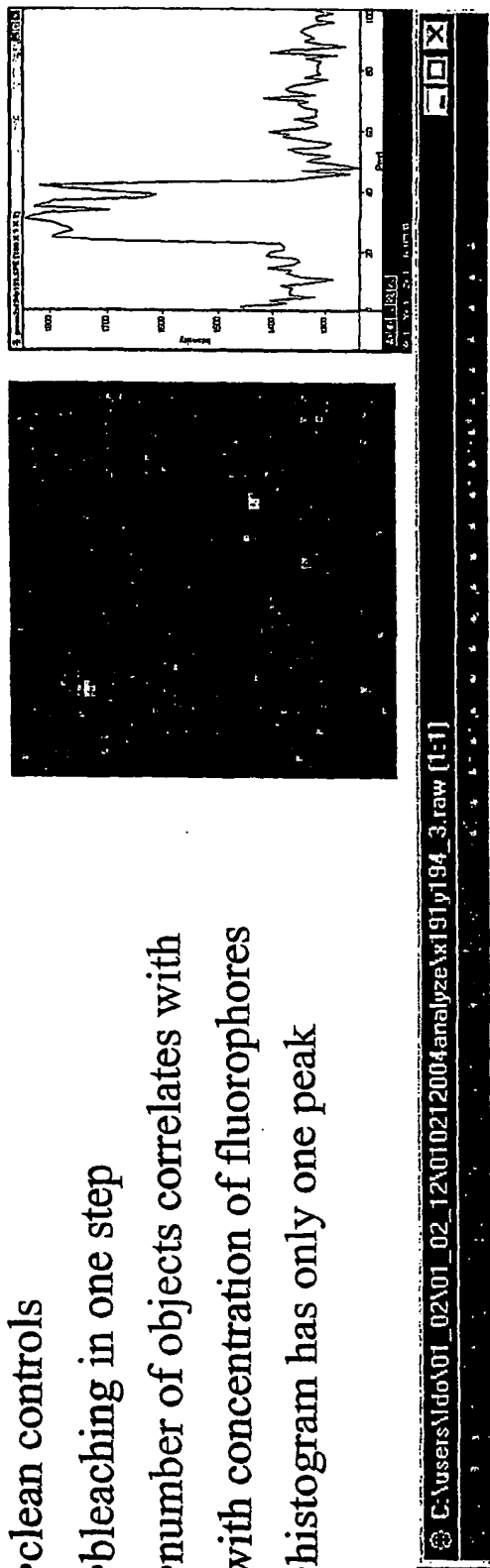


FIG. 6

# Bleaching in Steps Proves Multiple Incorporation

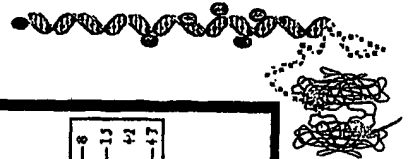
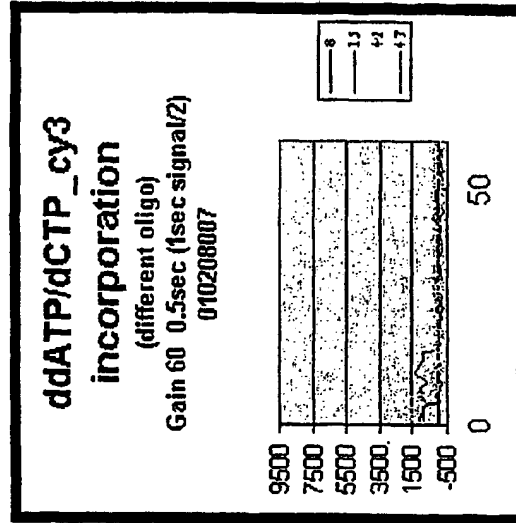
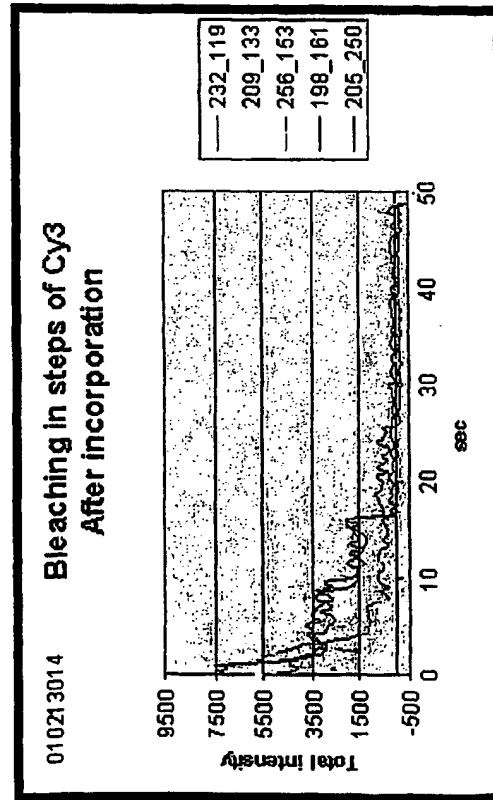
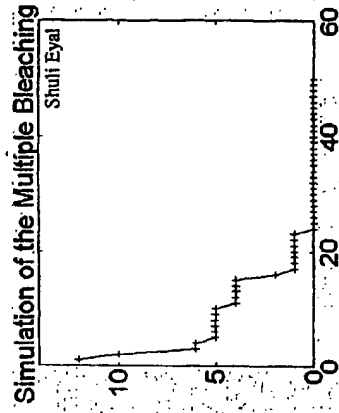
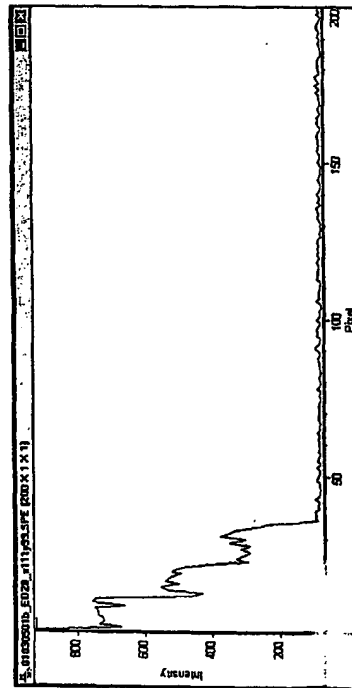
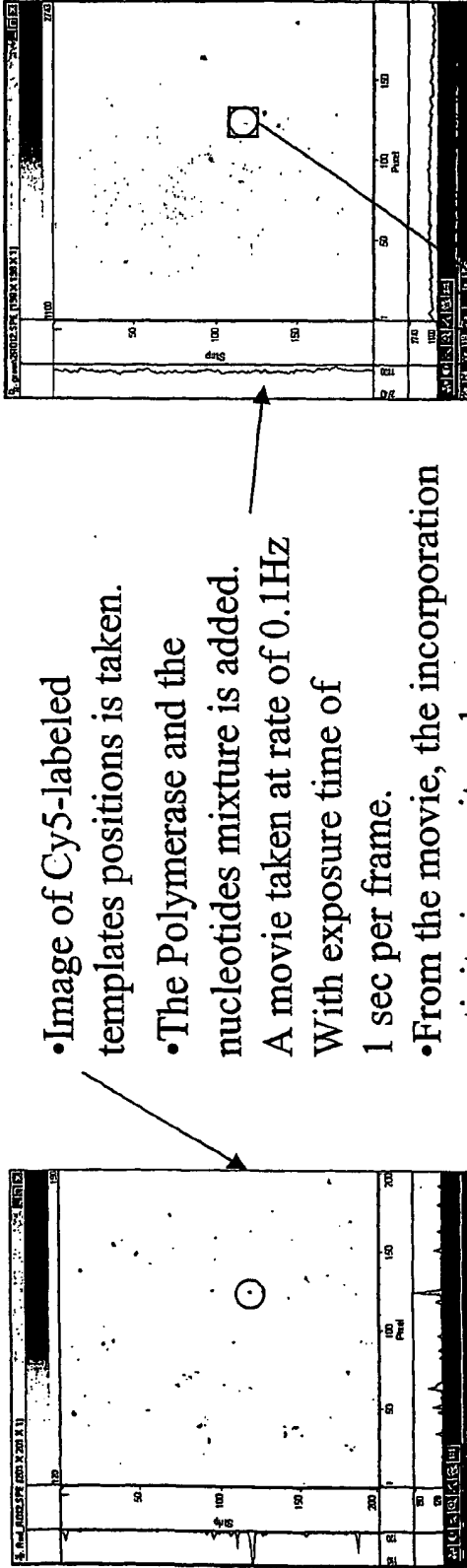


FIG. 7

# 1) Incorporation Dynamics

- Image of Cy5-labeled templates positions is taken.
  - The Polymerase and the nucleotides mixture is added.
- A movie taken at rate of 0.1Hz  
With exposure time of 1 sec per frame.
- From the movie, the incorporation activity is monitored



5 sec total exposure time



50 sec interval

FIG. 8

Interval between the frames is 10 sec.

Exposure time is 1 sec/frame.

# Incorporation events as a function of time

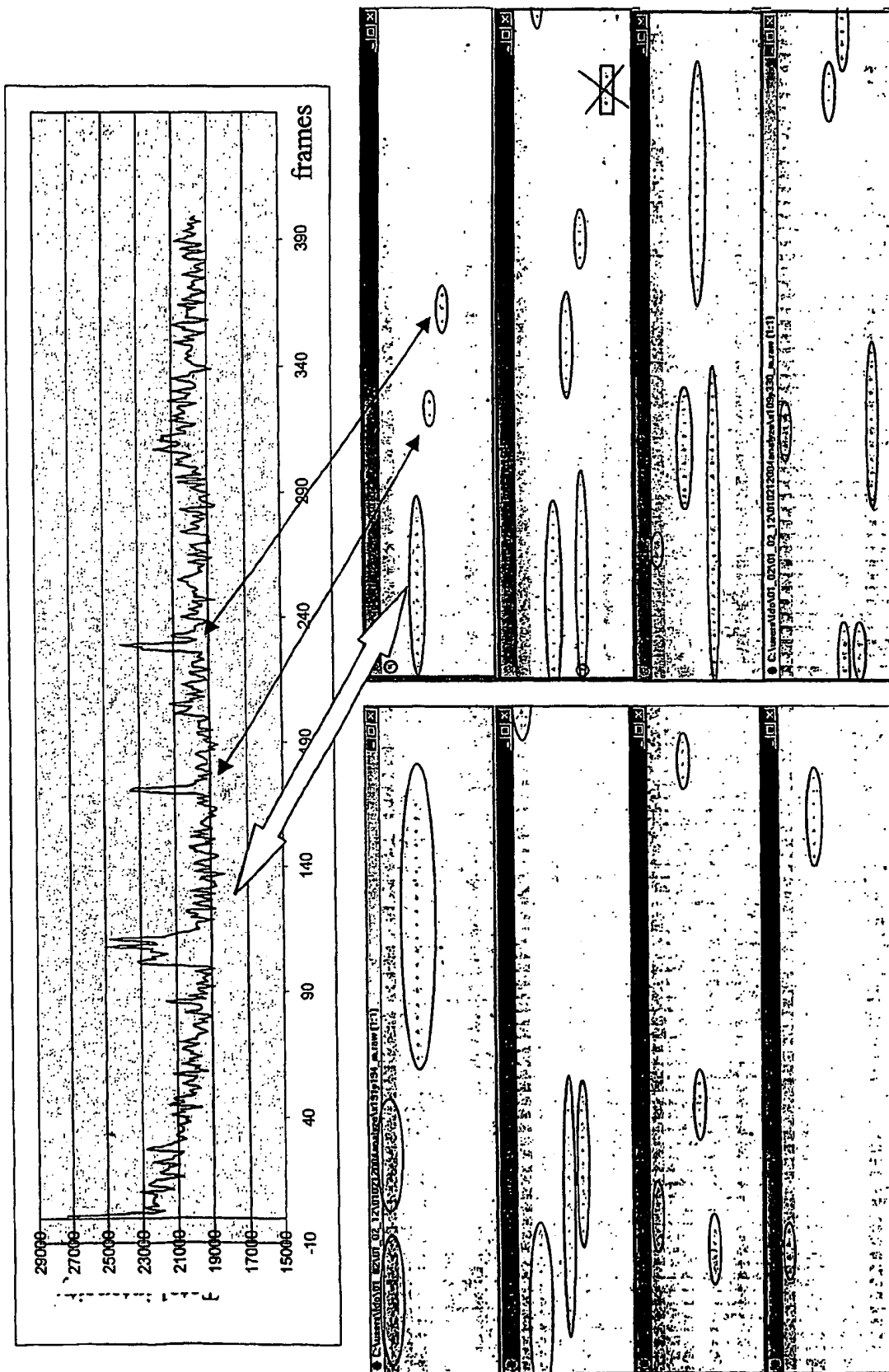


FIG. 9



# Statistics of Incorporation

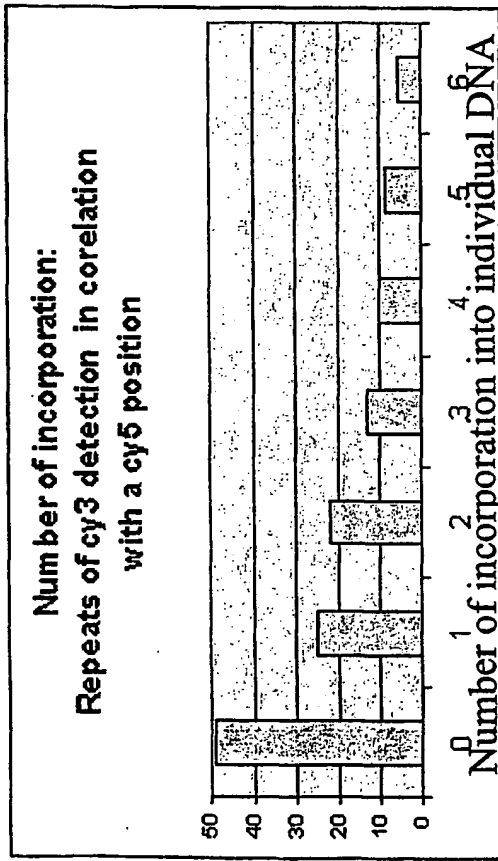
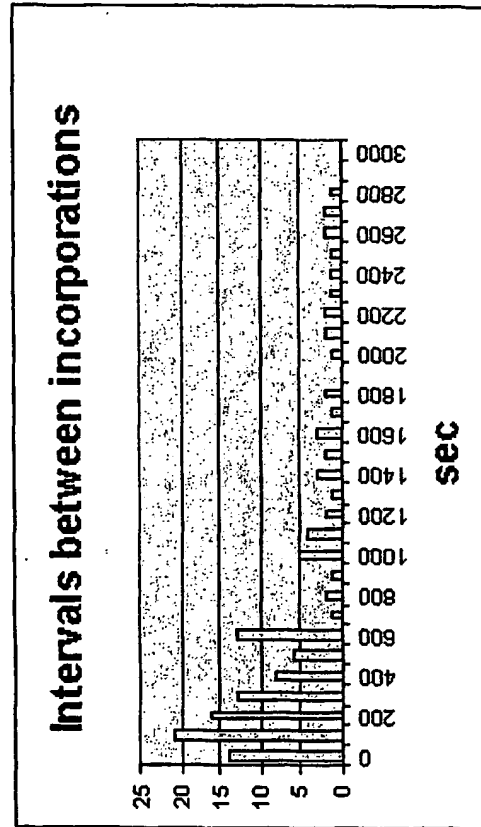
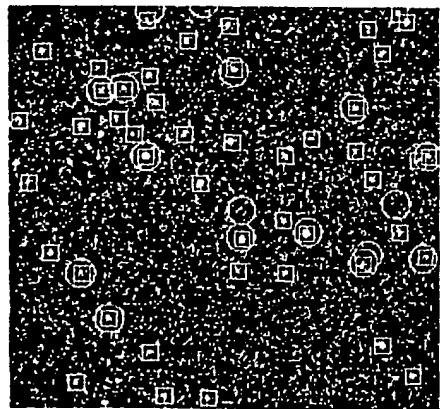


Fig. 10

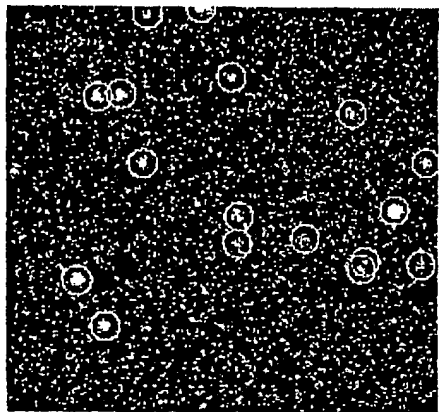


# Finding Correlation Between DNA Primer

## Location and Polymerase activity location



primer and incorporation features



Incorporation features

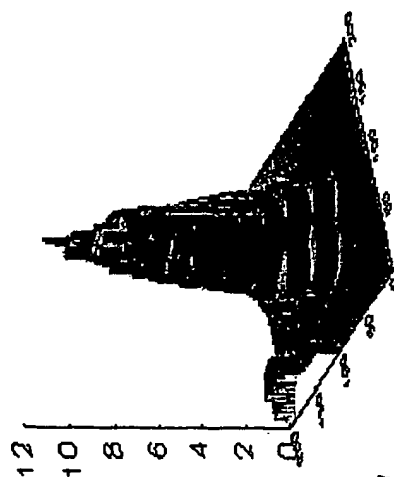
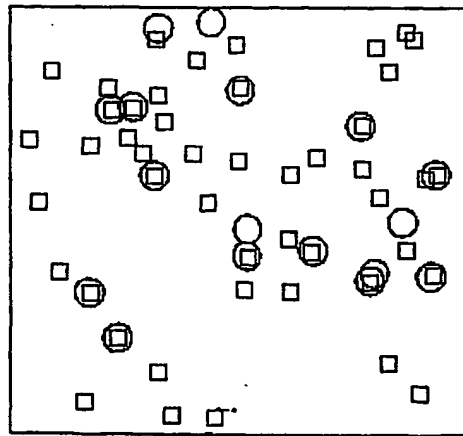


FIG. 11

# Determining DNA bases using correlation

DNA Sequence: 3'...ttttttA ccccccG...5'  
...aaaaaU? gggggGC

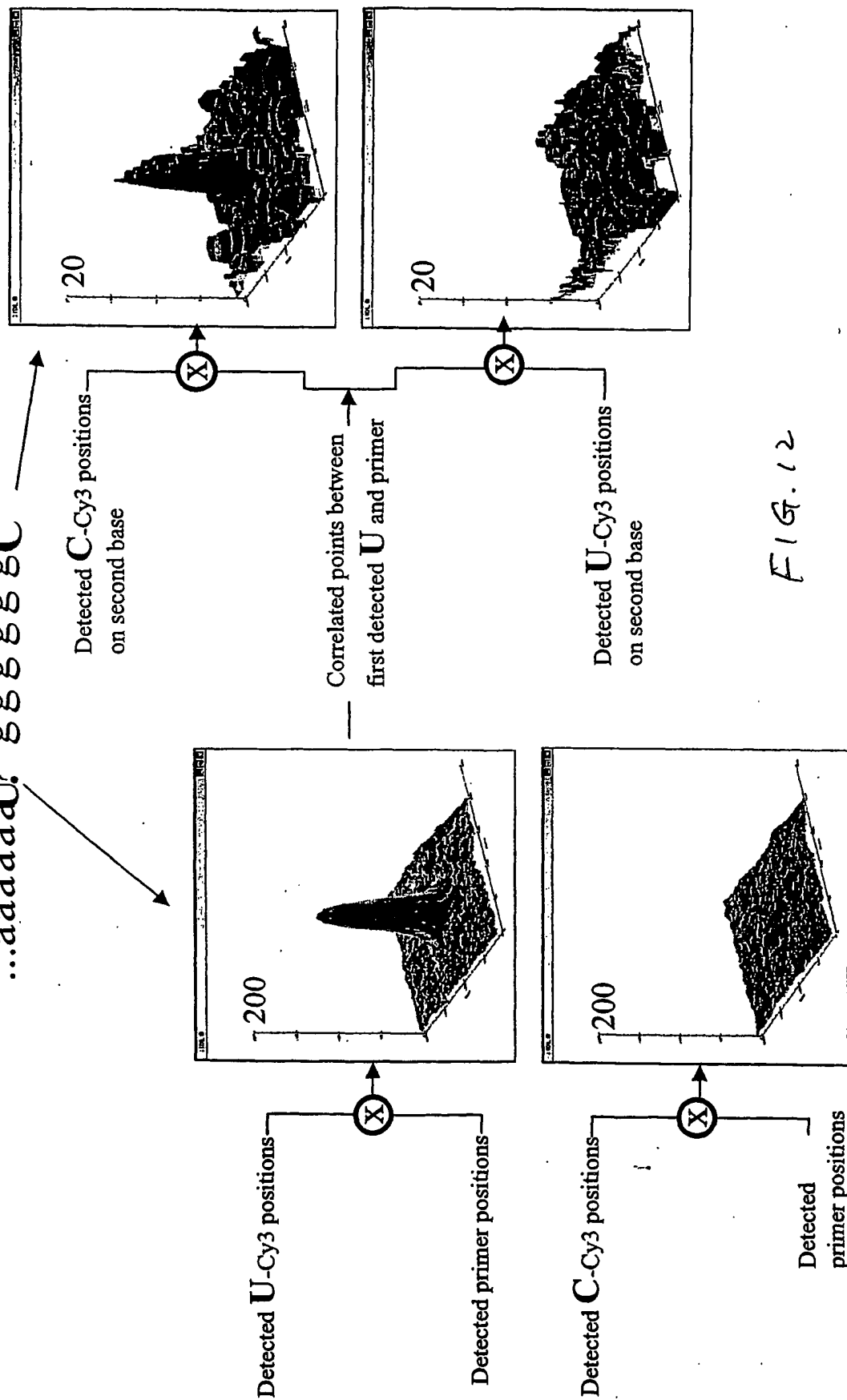


FIG. 12

# Detection of the Polymerase activity by FRET signals from Single DNA Molecules

- FRET signal after incorporation  
labeled nucleotide

- Fluorescently labeled DNA

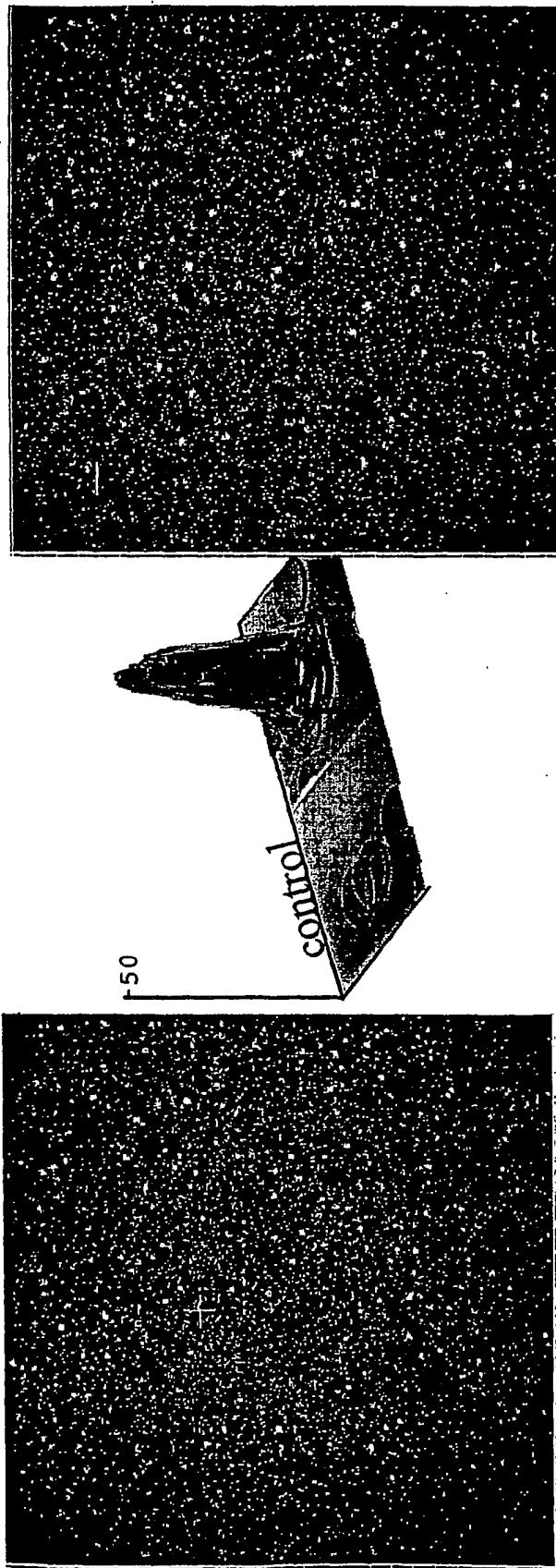


FIG. 13

..AttctttGcttcttAttctttGcttcttAttctttG.....

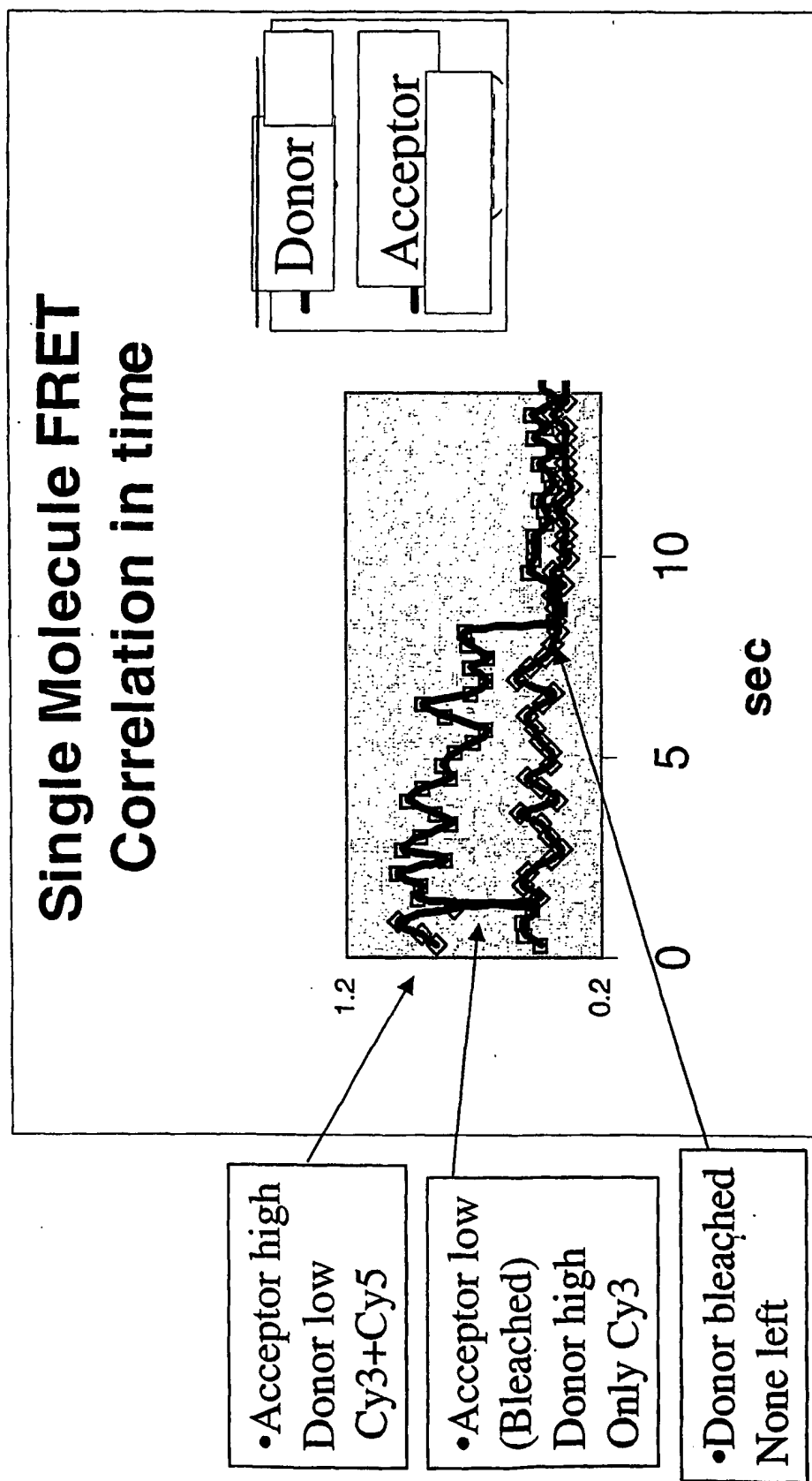
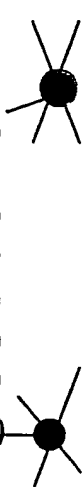


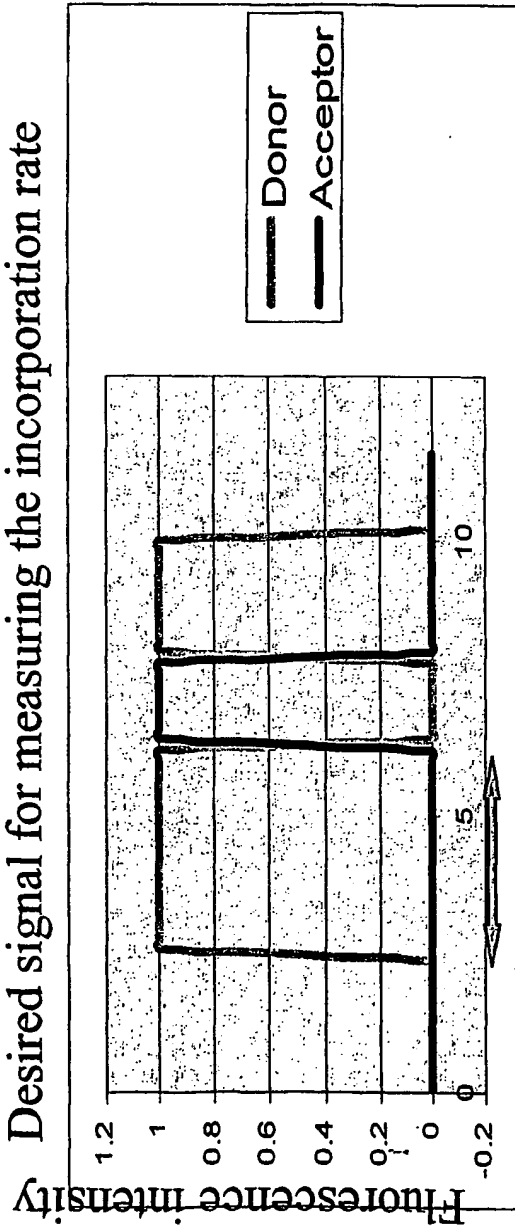
FIG. 14

# Studies of Polymerase Activity Using FRET

DNA Sequence: 3' ...tttttttA ccccccG...5'  
 ...aaaaaU ggggggGC



Desired signal for measuring the incorporation rate



time

Fig. 15

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/08187

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/68; C07H 21/02, 21/04

US CL : 435/6; 536/23.1, 24.3

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6; 536/23.1, 24.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
Please See Continuation Sheet

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,302,509 A (CHEESEMAN) 12 April 1994, see entire document.	1, 3-5, 22, 27
Y		2, 6-21, 23-26
Y, P	US 6,255,083 B1 (WILLIAMS) 03 July 2001, see entire document.	6, 17-19, 21 and 26
Y	US 5,832,165 A (REICHERT et al) 03 November 1998, see entire document.	6, 17-19, 21 and 26
Y	US 6,020,457 A (KLIMASH et al) 01 February 2000, see entire document.	9-15
A,P	US 6,221,592 B1 (SCHWARTZ et al) 24 April 2001, see entire document.	1-27

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

### Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"B" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

26 June 2002 (26.06.2002)

Date of mailing of the international search report

09 AUG 2002

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703)305-3230

Authorized officer

Brian Whisenant, Ph.D.

Telephone No. (703) 308-0196

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/08187

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claim Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claim Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:  
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-27

Remark on Protest

☐  
☐

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/08187

## BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claim(s) 1-27, drawn to a method of analyzing the sequence of a target ppolynucleotide.

Group II, claim(s) 28, drawn to an apparatus for analyzing the sequence of a target ppolynucleotide.

The inventions listed as Groups I and II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The claims as drawn are related to each other because they both relate to analyzing the sequence of a target polynucleotide. However, since the method of analyzing the sequence of a target polynucleotide, as claimed, is known, see for example, Cheeseman [US 5,302,509 (1994)], the claims are no longer linked by a special technical feature, because, by definition, the a special technical feature must distinguish over the prior art. Without the special technical feature the claims lack unity.

## Continuation of B. FIELDS SEARCHED Item 3:

USPATFULL via EAST, Medline, CAplus

search terms : sequencing, primer extension, polynucleotide or nucleic acid, fluor? and (TIRF ot total internal reflection fluorescence)

VIII-4-1	<b>Declaration: Inventorship (only for the purposes of the designation of the United States of America)</b> Declaration of Inventorship (Rules 4.17(iv) and 51bis.1(a)(iv)) for the purposes of the designation of the United States of America:	<p>I hereby declare that I believe I am the original, first and sole (if only one inventor is listed below), or joint (if more than one inventor is listed below) inventor of the subject matter which is claimed and for which a patent is sought.</p> <p>This declaration is directed to the international application of which it forms a part (if filing declaration with application).</p> <p>I hereby declare that my residence, mailing address, and citizenship are as stated next to my name.</p> <p>I hereby state that I have reviewed and understand the contents of the above-identified international application, including the claims of said application. I have identified in the request of said application, in compliance with PCT Rule 4.10, any claim to foreign priority, and I have identified below, under the heading "Prior Applications," by application number, country or Member of the World Trade Organization, day, month and year of filing, any application for a patent or inventor's certificate filed in a country other than the United States of America, including any PCT international application designating at least one country other than the United States of America, having a filing date before that of the application on which foreign priority is claimed.</p>
VIII-4-1 -1	Prior applications:	60/275,232, US, 12 March 2001 (12.03.2001)

CORRECTED VERSION

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
19 September 2002 (19.09.2002)

PCT

(10) International Publication Number  
**WO 2002/072892 A1**

(51) International Patent Classification<sup>7</sup>: **C12Q 1/68**,  
C07H 21/02, 21/04

(21) International Application Number:  
PCT/US2002/008187

(22) International Filing Date: 12 March 2002 (12.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/275,232 12 March 2001 (12.03.2001) US

(71) Applicant (for all designated States except US): CALIFORNIA INSTITUTE OF TECHNOLOGY [US/US]; 1200 East California Boulevard, MC 201-85, Pasadena, CA 91125 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **QUAKE**, Stephen [US/US]; 744 Plymouth Road, San Marino, CA 91108 (US). **BRASLAVSKY**, Ido [IL/US]; 120 South Chester Avenue, #7, Pasadena, CA 91106 (US). **HEBERT**, Benedict [CA/US]; 1028 East Del Mar Boulevard, #306, Pasadena, CA 91106 (US). **KARTALOV**, Emil [BG/US]; MC5027, Avery House, 1200 East California Boulevard, Pasadena, CA 91125 (US).

(74) Agents: **WANG**, Hugh et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111-3834 (US).

(81) Designated States (national): AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA,

CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK (utility model), SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declarations under Rule 4.17:**

- of inventorship (Rule 4.17(iv)) for US only
- of inventorship (Rule 4.17(iv)) for US only
- of inventorship (Rule 4.17(iv)) for US only
- of inventorship (Rule 4.17(iv)) for US only

**Published:**

- with international search report

(48) Date of publication of this corrected version:  
27 May 2004

(15) Information about Correction:  
see PCT Gazette No. 22/2004 of 27 May 2004, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND APPARATUS FOR ANALYZING POLYNUCLEOTIDE SEQUENCES BY ASYNCHRONOUS BASE EXTENSION

(57) Abstract: The invention provides methods and apparatus for analyzing polynucleotide sequences by asynchronous base extension. Some applications of the invention utilize total internal reflection fluorescence microscopy to image polynucleotide molecules at single molecule resolution.



WO 2002/072892 A1

# METHODS AND APPARATUS FOR ANALYZING POLYNUCLEOTIDE SEQUENCES BY ASYNCHRONOUS BASE EXTENSION

## CROSS-REFERENCES TO RELATED APPLICATIONS

5 This nonprovisional patent application claims the benefit of U.S. Provisional Patent Application No. 60/275,232, filed March 12, 2001, the disclosure of which is hereby incorporated by reference in its entirety and for all purposes.

## TECHNICAL FIELD

10 The present invention relates to novel methods and apparatus for analyzing polynucleotide sequences with high sensitivity and parallelism.

## BACKGROUND OF THE INVENTION

Methods for analyzing polynucleotide sequences can be grouped to two major  
15 fields: electrophoretic and non-electrophoretic methods. The electrophoretic methods include slab gel electrophoresis, capillary electrophoresis, microfabricated capillary arrays, and free solution electrophoresis. All these methods rely on the Sanger method in which polynucleotide chain elongation inhibitors are incorporated into the polynucleotide strands which are then separated according to their sizes, usually on a polyacrylamide gel. These  
20 methods are the common means for analyzing polynucleotide sequences nowadays. However, the process is time-consuming, requires large amount of target polynucleotides and reaction reagents, and has limited ability to read long sequences that are inherent in the gel electrophoresis methods. The non-electrophoretic methods include pyrosequencing, sequencing by hybridization, massively parallel signature sequencing, and sequencing by  
25 mass spectrometry. These methods also have a number of disadvantages. For example, they usually require synchronization of the polynucleotide templates which inevitably decay with each cycle of sequencing reaction.

Thus, there is a need in the art for better methods for analyzing polynucleotide sequences, e.g., methods with high throughput, parallelism, and resolution. The present  
30 invention fulfills this and other needs.

## SUMMARY OF THE INVENTION

In one aspect, the present invention provides methods for analyzing the sequence of a target polynucleotide. The methods include the steps of (a) providing a primed target polynucleotide immobilized to a surface of a substrate; wherein the target  
5 polynucleotide is attached to the surface with single molecule resolution; (b) In the presence of a polymerase, adding a first fluorescently labeled nucleotide to the surface of the substrate under conditions whereby the first nucleotide attaches to the primer, if a complementary nucleotide is present to serve as template in the target polynucleotide; (c) determining presence or absence of a fluorescence signal on the surface where the target polynucleotide is  
10 immobilized, the presence of a signal indicating that the first nucleotide was incorporated into the primer, and hence the identity of the complementary base that served as a template in the target polynucleotide; and (d) repeating steps (b)-(c) with a further fluorescently labeled nucleotide, the same or different from the first nucleotide, whereby the further nucleotide attaches to the primer or a nucleotide previously incorporated into the primer.

15 In some methods, a plurality of different primed target polynucleotides are immobilized to different portions of the substrate. In some methods, steps (b)-(c) are performed at least four times with four different types of labeled nucleotides. In some methods, steps (b)-(c) are performed until the identity of each base in the target polynucleotide has been identified. In some methods, there is an additional step of removing  
20 the signal after step (c). In some methods, all ingredients are present simultaneously and a continues monitoring of the incorporation is facilitated.

In some methods of the invention, the presence or absence of a fluorescence signal is determined with total internal reflection fluorescence (TIRF) microscopy. In some methods, the target polynucleotide is primed with a fluorescently labeled primer (e.g., with  
25 Cy5 or Cy3). Some methods of the invention employ nucleotides that are labeled with Cy3 or Cy5.

Various materials can be used to immobilize the target polynucleotides. In some methods, a fused silica or glass slide is used. In some methods, the substrate surface is coated with a polyelectrolyte multilayer (PEM). The PEM can be terminated with a  
30 polyanion, which helps to repel nucleotides from the surface and reduce non-specific binding to the surface. The polyanion can bear pendant carboxylic acid groups. In some of these methods, the target polynucleotide is biotinylated, and the substrate surface is coated with streptavidin. Often the surface is coated with biotin prior to coating with streptavidin. In

some methods, the surface is coated with a polyelectrolyte multilayer (PEM) terminated with carboxylic acid groups prior to attachment of biotin.

In some methods of the invention, a light source for illuminating the surface of said substrate and a detection system for detecting a signal from said surface are employed.

- 5    Optionally, an appropriately programmed computer is also employed for recording identity of a nucleotide when the nucleotide becomes incorporated into the immobilized primer.

In another aspect, the invention provides apparatus for carrying out the methods of the invention. Typically, the apparatus contain (a) a flow cell which houses a substrate for immobilizing target polynucleotide(s) with single molecule resolution; (b) an  
10    inlet port and an outlet port in fluid communication with the flow cell for flowing fluids into and through the flow cell; (c) a light source for illuminating the surface of the substrate; and (d) a detection system for detecting a signal from said surface. Some of the apparatus are microfabricated. In some of these apparatus, the substrate is a microfabricated synthesis channel.

- 15    A further understanding of the nature and advantages of the present invention may be realized by reference to the remaining portions of the specification, the figures and claims.

All publications, patents, and patent applications cited herein are hereby expressly incorporated by reference in their entirety and for all purposes to the same extent as  
20    if each was so individually denoted.

### BRIEF DESCRIPTION OF THE DRAWINGS

- Figure 1 shows schematically immobilization of a primed polynucleotide and  
25    incorporation of labeled nucleotides.

Figure 2 shows schematically the optical setup of a detection system for total internal reflection microscopy.

- 30    Figure 3 shows results which indicate that streptavidin is required for immobilizing the polynucleotide template in an exemplified embodiment.

Figure 4 shows results which indicate that DNA polymerase incorporating labeled nucleotide into the immobilized primer is visualized with single molecule resolution.

Figure 5 shows incorporation of multiple labeled nucleotides in a bulk experiment in solution, using biotin-labeled 7G oligonucleotide template (SEQ ID NO:1) and p7G primer (SEQ ID NO:2).

5

Figure 6 shows low background signal from free nucleotides in solution and detection of signals from incorporated nucleotides.

Figure 7 shows results from experiments and simulation of multiple bleaching.

10

Figure 8 shows dynamics of incorporation of labeled nucleotides into the immobilized primer.

Figure 9 shows multiple incorporation events of labeled nucleotides over a period of time.

15

Figure 10 shows statistics of incorporation of labeled nucleotides over a period of time.

Figure 11 shows correlation between location of labeled primer and location of incorporation of labeled nucleotides.

20

Figure 12 shows correlation graphs for incorporation of two labeled nucleotides, using a 6TA6GC oligonucleotide template (SEQ ID NO:6) and a p7G primer (SEQ ID NO:2). Partial sequences of the template, 5' - GccccccAttttt - 3' (SEQ ID NO:7), and the extended product, 5' - aaaaaaUggggggC (SEQ ID NO:8), are also shown in the Figure.

25

Figure 13 shows detection of fluorescence resonance energy transfer (FRET) when two different labels are incorporated into the same primer. The polynucleotide template used here is the 7G7A oligonucleotide (SEQ ID NO:5), but only part of the sequence, 5' - AttctttGcttcttAttctttGcttcttAttctttG - 3' (SEQ ID NO:9), is shown in the Figure.

30

Figure 14 shows correlation of single molecule FRET signals over a period of time.

Figure 15 shows the expected signals from an experiment in which two colors, donor and acceptor, are incorporated one after the another. Partial sequences of the template, 5' - GccccccAttttt - 3' (SEQ ID NO:7), and the extended product, 5' - aaaaaaUggggggC (SEQ ID NO:8), are also shown in the Figure.

## DETAILED DESCRIPTION

### I. Overview

The present invention provides methods and apparatus for analyzing polynucleotides with high sensitivity, parallelism, and long read frames. The invention is predicated in part on visualization of incorporation of labeled nucleotides into immobilized polynucleotide template molecules in a time resolved manner with single molecule resolution. As each of the immobilized template molecules is read individually, no synchronization is needed between the different molecules. Instead, with methods of the present invention, asynchronous base extension is sufficient for analyzing a target polynucleotide sequence.

In some aspects of the invention, single molecule resolution was achieved by immobilizing the template molecules at very low concentration to a surface of a substrate, coating the surface to create surface chemistry that facilitates template attachment and reduces background noise, and imaging nucleotide incorporation with total internal reflection fluorescence microscopy. Analysis with single molecule resolution provides the advantage of monitoring the individual properties of different molecules. It allows identification of properties of an individual molecule that can not be revealed by bulk measurements in which a large number of molecules are measured together. Furthermore, to determine kinetics, bulk measurements require synchronization of the molecules or system state, while in single molecule analysis there is no need for synchronization.

The polynucleotides suitable for analysis with the invention can be DNA or RNA. The analysis can be for sequence analysis, DNA fingerprinting, polymorphism identification, or gene expression measurement. The methods can also be used to analyze activities of other biomacromolecules such as RNA translation and protein assembly. In a preferred embodiment, the method entails immobilization of primed polynucleotide templates to the surface of a solid substrate (e.g., a glass slide). The templates are pre-hybridized to a



labeled primer (e.g., with a fluorescent dye) so that their location on the surface can be imaged with single molecule sensitivity. An evanescent light field is set up at the surface in order to image the fluorescently labeled polynucleotide molecules. The evanescent field is also used to image fluorescently labeled nucleotide triphosphates (dNTPs or NTPs) upon  
5 their incorporation into the immobilized primer when a polymerase is present.

Methods of the present invention find various applications in polynucleotide sequence analysis. In some applications, a static approach is employed. Such an approach involves adding just one type of labeled nucleotide to the extension reaction at any given time. The signal is incorporated into the primer if the next template residue in the target  
10 polynucleotide is the complementary type. Otherwise, a different type of labeled nucleotide is used until the correct residue is incorporated. In other applications, a dynamic approach is employed. In these methods, all four types of nucleotides (at least one type labeled) are simultaneously present in the reaction, and incorporation of the signals into the primer is monitored dynamically. For example, incorporated signals are imaged continuously,  
15 preferably at a rate faster than the rate at which the nucleotides are incorporated into the primer.

Preferably, visualization of the templates or incorporated nucleotides are realized with total internal reflection (TIR) fluorescence microscopy. With TIR technology, the excitation light (e.g., a laser beam) illuminates only a small volume of liquid close to the  
20 substrate (excitation zone). Signals from free nucleotides in solution that are not present in the excitation zone are not detected. Signals from free nucleotides that diffuse into the excitation zone appear as a broad band background because the free nucleotides move quickly across the excitation zone. Optionally, the fluorescence signals are removed by photobleaching or by chemical means after one or more rounds of incorporation. The  
25 methods can also employ microfluidic means to control flow of reaction reagents. In such methods, labeled nucleotides and other reaction reagents can be exchanged in a fast and economic way.

Further, employing a microfluidic device which allows fast fluid exchange, concentrations of nucleotides and/or other reaction reagents can be alternated at different time  
30 points of the analysis. This could lead to increased incorporation rate and sensitivity of the analysis. For example, when all four types of nucleotides are simultaneously present in the reaction to monitor dynamic incorporation of nucleotides, concentrations of the nucleotides can be alternated between  $\mu\text{M}$  range and sub-nM range. This leads to both better visualization of the signals when low concentrations of nucleotides are present, and increased

polymerization rate when higher concentrations of nucleotides are present. Using a microfluidic device, the rate at which the concentrations can be alternated can be as high as a few tens of Hertz. Alternating concentrations of nucleotides is also beneficial to improving signal visualization and polymerization rate in the static approach of sequence analysis. In this approach, after adding a given type of labeled nucleotide to the immobilized template/primer complex and sufficient time for incorporation, free nucleotides (as well as other reaction reagents in solution) can be flown out using a microfluidic device. This will leave a much lower concentration of free nucleotide when the signals are visualized. Optionally, an additional washing step can be employed to further reduce the free nucleotide concentration before the signals are imaged.

In some methods, polynucleotide sequence analysis is accomplished by using four different fluorescent labels on the four nucleotide triphosphates. Incorporated signals are imaged and then photobleached before the next incorporation cycle. Runs of identical bases (e.g., AAAAA) can be identified by, e.g., monitoring the intensity of the signal so that the number of fluorophores at the emitting spot can be determined. Further, signals due to fluorescence resonance energy transfer (FRET) can be detected from individual DNA strands when two different type of fluorescent dyes are incorporated into the same DNA. Such signals are useful to determine sequence information of the immobilized template polynucleotide.

Thus, in some methods, multiple types of labeled nucleotides (e.g., 2 to 4 types each labeled with a different fluorescent dye) can be added at the same time for the extension reactions. In some methods, one type of labeled nucleotide is added at a step, and each extension cycle may comprise four such steps in order to observe the incorporation of a complementary nucleotide. In some methods, less than all four dNTPs are labeled. For example, the analysis can have only two of the nucleotides labeled. By repeating the experiment with different pairs (e.g., AT, AG, AC, TG, TC, GC), the original nucleotide sequence can be delineated. In some methods, the incorporation/extension reaction is performed with multiple copies of the template polynucleotide. Alternatively, one immobilized template molecule can be used repeatedly, by denaturing the extended molecule, removing the newly synthesized strand, annealing a new primer, and then repeating the experiment in situ with fresh reagents.

The present invention is also useful to obtain partial sequence information of a target polynucleotide, e.g., by using only two or three labeled nucleotide species. The relative positions of two or three nucleotide species in the sequence in conjunction with

known sequence databases can facilitate determination of the identity of the target sequence, i.e., whether it is identical or related to a known sequence. Such an approach is useful, for example, in determining gene expressions by sequencing cDNA libraries.

The present methods avoid many of the problems observed with the prior art sequencing methods. For example, the methods are highly parallel since many molecules are analyzed simultaneously and in high density (e.g., one template molecule per  $\sim 10\mu\text{m}^2$  of surface area). Thus, many different polynucleotides can be sequenced or genotyped on a single substrate surface simultaneously. In addition, stepwise addition of nucleotides is unnecessary in some methods, as all four nucleotides can be added simultaneously. Rather, sequence information is produced continuously as polymerases continually incorporate all four nucleotides into growing polynucleotide chains. The methods are also extremely sensitive because information obtained from only a single copy of the template molecule is needed in order to determine its sequence. Releasing the extension product from the polynucleotide template, e.g., by denaturing and annealing the template with a different primer provides the opportunity to read again the same template molecule with different sets of nucleotides (e.g., different combinations of two types of labeled nucleotide and two types of unlabeled nucleotides).

## II. Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which this invention pertains. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton *et al.*, DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY (2d ed. 1994); THE CAMBRIDGE DICTIONARY OF SCIENCE AND TECHNOLOGY (Walker ed., 1988); and Hale & Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY (1991). Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described. The following definitions are provided to assist the reader in the practice of the invention.

"Array" refers to a solid support having more than one site or location having either a target polynucleotide or a polymerase bound thereto.

A "base" or "base-type" refers to a particular type of nucleoside base. Typical bases include adenine, cytosine, guanine, uracil, or thymine bases where the type refers to the subpopulation of nucleotides having that base within a population of nucleotide triphosphates

bearing different bases. Other rarer bases or analogs can be substituted such as xanthine or hypoxanthine or methylated cytosine.

“Complements a region of the target nucleic acid downstream of the region to be sequenced” in the context of sequencing or genotyping refers to the fact that the primers are extended in a 3' direction by a polymerase. Therefore the primer binds to a subsequence of the target 3' (downstream) to the target sequence that is to be determined as the 3' end of the primer is extended.

“Genotyping” is a determination of allelic content of a target polynucleotide without necessarily determining the sequence content of the entire polynucleotide. It is a subset of sequencing. For example the identification of single nucleotide polymorphisms by determination of single base differences between two known forms of an allele is a form of sequencing that does not require all the target polynucleotide to be sequenced.

“Immobilizing” refers to the attachment of a target nucleic acid or polymerase to a solid support by a means that prevents its release in a reaction solution. The means can be covalent bonding or ionic bonding or hydrophobic bonding.

“Nucleoside” includes natural nucleosides, including ribonucleosides and 2'-deoxyribonucleosides, as well as nucleoside analogs having modified bases or sugar backbones.

The terms “nucleic acid” or “nucleic acid molecule” refer to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, can encompass known analogs of natural nucleotides that can function in a similar manner as naturally occurring nucleotides. Unless otherwise noted, “nucleic acid” and “polynucleotide” are used interchangeably.

“Oligonucleotide” or “polynucleotide” refers to a molecule comprised of a plurality of deoxyribonucleotides or nucleoside subunits. The linkage between the nucleoside subunits can be provided by phosphates, phosphonates, phosphoramidates, phosphorothioates, or the like, or by nonphosphate groups as are known in the art, such as peptide-type linkages utilized in peptide nucleic acids (PNAs). The linking groups can be chiral or achiral. The oligonucleotides or polynucleotides can range in length from 2 nucleoside subunits to hundreds or thousands of nucleoside subunits. While oligonucleotides are preferably 5 to 100 subunits in length, and more preferably, 5 to 60 subunits in length, the length of polynucleotides can be much greater (e.g., up to 100 kb). (...if a whole chromosome is targeted...Thought 100kb will be already nice..) [“e.g.” means it is not exclusive. Also, “100 Mb” probably does not make practical sense]

"Optical reader" or "detection system" refers to a device that can detect and record light emitted from the labeled dNTP (or NTP) or immobilized polynucleotide template (and/or primer) molecules.

The term "primer" refers to an oligonucleotide, whether occurring naturally as  
5 in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, (i.e., in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature, buffer and pH). The primer is preferably single stranded for maximum  
10 efficiency in amplification, but can alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. Preferably, the primer is an oligodeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of the primers depend on many factors, including temperature, source of primer and  
15 the use of the method.

A primer is selected to be "substantially" complementary to a strand of specific sequence of the template. A primer must be sufficiently complementary to hybridize with a template strand for primer elongation to occur. A primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment can  
20 be attached to the 5' end of the primer, with the remainder of the primer sequence being substantially complementary to the strand. Non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementarity with the sequence of the template to hybridize and thereby form a template primer complex for synthesis of the extension product of the primer. The use of random  
25 primer is used in some cases. For example, when the terminal sequence of the target or template polynucleotide is not known, random primer combinations can be used.

The term "probe" refers to an oligonucleotide (i.e., a sequence of nucleotides), whether occurring naturally as in a purified restriction digest or produced synthetically, recombinantly or by PCR amplification, which is capable of hybridizing to another  
30 oligonucleotide of interest. A probe can be single-stranded or double-stranded. Probes are useful in the detection, identification and isolation of particular gene sequences. It is contemplated that any probe used in the present invention can be labeled with any "reporter molecule," so that is detectable in any detection system, including, but not limited to fluorescent, enzyme (e.g., ELISA, as well as enzyme-based histochemical assays),

radioactive, quantum dots, and luminescent systems. It is not intended that the present invention be limited to any particular detection system or label.

"Sequencing" refers to the determination of the order and position of bases in a polynucleotide molecule.

5 "Single molecule configuration" refers to an array of molecules on a solid support where members of the array are present as an individual molecule located in a defined location. The members can be the same or different.

"Single molecule resolution" refers to the ability of a system to resolve one molecule from another. For example, in far field optical system the detection limit is in the  
10 order of a micron. This implies that the distance between two identical molecules to be resolved is at least few microns apart.

"Specific hybridization" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions. Stringent conditions are conditions under which a probe can hybridize to its target subsequence, but to  
15 no other sequences. Stringent conditions are sequence-dependent and are different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5° C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50%  
20 of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Typically, stringent conditions include a salt concentration of at least about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides). Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide or tetraalkyl ammonium  
25 salts. For example, conditions of 5X SSPE (750 mM NaCl, 50 mM Na Phosphate, 5 mM EDTA, pH 7.4) and a temperature of 25-30°C are suitable for allele-specific probe hybridizations. (See Sambrook *et al.*, *Molecular Cloning* 2001).

The term "template" or "target" refers to a polynucleotide of which the sequence is to be analyzed. In some cases "template" is sought to be sorted out from other  
30 polynucleotide sequences. "Substantially single-stranded template" is polynucleotide that is either completely single-stranded (having no double-stranded areas) or single-stranded except for a proportionately small area of double-stranded polynucleotide (such as the area defined by a hybridized primer or the area defined by intramolecular bonding). "Substantially double-stranded template" is polynucleotide that is either completely double-stranded (having

no single-stranded region) or double-stranded except for a proportionately small area of single-stranded polynucleotide.

### III. Template Preparation and Immobilization

#### 5 A. Introduction

This invention provides novel methods and apparatus to analyze polynucleotide sequences (e.g., sequencing and genotyping). Preferably, the target or template polynucleotide to be analyzed is immobilized to the surface of a solid substrate (e.g., a fused silica slide) at single molecule resolution. Preferably, the polynucleotide is pre-  
10 hybridized to a labeled primer. A DNA or RNA polymerase, four different types of nucleotide triphosphates (NTPs or dNTPs, depending on the template and polymerase used), and other reaction reagents are then applied to the immobilized polynucleotide. At least one type of the nucleotides are fluorescently labeled. When more than one type of NTPs are labeled, the labels are preferably different for different NTPs. Using TIR fluorescent  
15 microscopy, incorporation of the labeled nucleotide into a target or template polynucleotide is detected by imaging fluorescence signal from the immobilized polynucleotide with single molecule resolution. Preferably, all four labeled NTPs are present simultaneously. As the polymerase continues to move along the target polynucleotide, the polynucleotide sequence is read from the order of the incorporated labels.

#### 20 B. Target or template polynucleotide

The target polynucleotide is not critical and can come from a variety of standard sources. It can be mRNA, ribosomal RNA, genomic DNA or cDNA. They can comprise naturally occurring and or non-naturally occurring nucleotides. Templates suitable for analysis according to the present invention can have various sizes. For example, the  
25 template can have a length of 100 bp, 200 bp, 500 bp, 1 kb, 3 kb, 10 kb, or 20 kb and so on. When the target is from a biological source, there are a variety of known procedures for extracting polynucleotide and optionally amplified to a concentration convenient for genotyping or sequence work. Polynucleotide can be obtained from any living cell of a person, animal or plant. Humans, pathogenic microbes and viruses are particularly  
30 interesting sources.

Polynucleotide amplification methods are known in the art. Preferably, the amplification is carried out by polymerase chain reaction (PCR). See, U.S. Pat. Nos. 4,683,202, 4,683,195 and 4,889,818; Gyllenstein et al., 1988, Proc. Natl. Acad. Sci. USA 85:

7652-7656; Ochman et al., 1988, Genetics 120: 621-623; Loh et al., 1989, Science 243: 217-220; Innis et al., 1990, PCR Protocols, Academic Press, Inc., San Diego, Calif. Other amplification methods known in the art that can be used in the present invention include ligase chain reaction (see EP 320,308), or methods disclosed in Kricka et al., 1995,

5 *Molecular Probing, Blotting, and Sequencing*, Chap. 1 and Table IX, Academic Press, New York.

### C. Primer annealing

Primers in combination with polymerases are used to sequence target  
10 polynucleotide. Primer length is selected to provide for hybridization to complementary template polynucleotide. The primers will generally be at least 10 bp in length, usually between 15 and 30 bp in length. If part of the template sequence is known, a specific primer can be constructed and hybridized to the template. Alternatively, if sequence of the template is completely unknown, the primers can bind to synthetic oligonucleotide adaptors joined to  
15 the ends of target polynucleotide by a ligase.

In some methods, the primer is labeled. When hybridized to the immobilized template, the labeled primer facilitates imaging location of the template. As exemplified in the Examples below, the primer can be labeled with a fluorescent label (e.g., Cy5). Preferably, the label used to label the primer is different from the labels on the nucleotides in  
20 the subsequent extension reactions.

The primers can be synthetically made using conventional nucleic acid synthesis technology. For example, the primers can be conveniently synthesized on an automated DNA synthesizer, e.g. an Applied Biosystems, Inc. (Foster City, Calif.) model 392 or 394 DNA/RNA Synthesizer, using standard chemistries, such as phosphoramidite  
25 chemistry, e.g. disclosed in the following references: Beaucage and Iyer, Tetrahedron, 48: 2223-2311 (1992); Molko et al, U.S. Pat. No. 4,980,460; Koster et al, U.S. Pat. No. 4,725,677; Caruthers et al, U.S. Pat. Nos. 4,415,732; 4,458,066; and 4,973,679; and the like. Alternative chemistries, e.g. resulting in non-natural backbone groups, such as phosphorothioate, phosphoramidate, and the like, may also be employed provided that the  
30 resulting oligonucleotides are compatible with the polymerase. The primers can also be ordered commercially from a variety of companies which specialize in custom oligonucleotides such as Operon Inc (Alameda, California).

Primer annealing is performed under conditions which are stringent enough to achieve sequence specificity yet sufficiently permissive to allow formation of stable hybrids



at an acceptable rate. The temperature and length of time required for primer annealing depend upon several factors including the base composition, length and concentration of the primer, and the nature of the solvent used, e.g., the concentration of DMSO, formamide, or glycerol, and counter ions such as magnesium. Typically, hybridization with synthetic polynucleotides is carried out at a temperature that is approximately 5 to 10°C below the melting temperature of the target-primer hybrid in the annealing solvent. In some methods, the annealing temperature is in the range of 55 to 75°C. and the primer concentration is approximately 0.2  $\mu$ M. Other conditions of primer annealing are provided in the Examples below. Under these preferred conditions, the annealing reaction can be complete in only a few seconds.

#### D. Immobilization of template polynucleotide

Preferably, the template or target polynucleotide molecules are provided as single molecule arrays immobilized to the surface of a solid substrate. The substrate can be glass, silica, plastic or any other conventionally non-reactive material that will not create significant noise or background for the fluorescent detection methods. Substrate surface to which the template polynucleotides are to be immobilized can also be the internal surface of a flow cell in a microfluidic apparatus, e.g., a microfabricated synthesis channel of the apparatus as described in the PCT application of Quake et al. (WO 01/32930; which is incorporated herein by reference). In some preferred embodiments, the solid support is made from fused silica slide (e.g., a fused silica glass slide from Esco, Cat. R130110). Compared to other support materials (e.g., a regular glass slide), fused silica has very low auto-fluorescence.

In some applications of the present invention, the template or target polynucleotides are immobilized to the substrate surface with single molecule resolution. In such methods, as exemplified in the Examples below, single molecule resolution is achieved by using very low concentration of the polynucleotide in the immobilization reaction. For example, a 10 pM concentration for a 80-mer polynucleotide template allows attachment of the polynucleotide to the surface of a silica slide at single molecule resolution (see Example 1). Template immobilization with single molecule resolution can also be verified by measuring bleach pattern of the fluorescently labeled templates (see Example 5).

In some methods, the templates are hybridized to the primers first and then immobilized to the surface. In some methods, the templates are immobilized to the surface prior to hybridization to the primer. In still some methods, the primers are immobilized to the

surface, and the templates are attached to the substrates through hybridization to the primers. In still some methods, the polymerase is immobilized to the surface.

Various methods can be used to immobilize the templates or the primers to the surface of the substrate. The immobilization can be achieved through direct or indirect

5 bonding of the templates to the surface. The bonding can be by covalent linkage. See, Joos et al., *Analytical Biochemistry* 247:96-101, 1997; Oroskar et al., *Clin. Chem* 42:1547-1555, 1996; and Khandjian, *Mole. Bio. Rep.* 11:107-115, 1986. The bonding can also be through non-covalent linkage. For example, Biotin-streptavidin (Taylor et al., *J. Phys. D. Appl. Phys.* 24:1443, 1991) and digoxigenin and anti-digoxigenin (Smith et al., *Science* 253: 1122, 1992) are common tools for attaching polynucleotides to surfaces and parallels. Alternatively, the bonding can be achieved by anchoring a hydrophobic chain into a lipidic monolayer or bilayer. When biotin-streptavidin linkage is used to immobilize the templates, the templates are biotinylated, and one surface of the substrates are coated with streptavidin. Since streptavidin is a tetramer, it has four biotin binding sites per molecule. Thus, it can provide linkage between the surface and the template. In order to coat a surface with streptavidin, the surface can be biotinylated first, and then parts of the four binding sites of streptavidin can be used to anchor the protein to the surface, leaving the other sites free to bind the biotinylated template (see, Taylor et al., *J. Phys. D. Appl. Phys.* 24:1443, 1991). Such treatment leads to a high density of streptavidin on the surface of the substrate, allowing a correspondingly high density of template coverage. Surface density of the template molecules can be controlled by adjusting concentration of the template which is applied to the surface. Reagents for biotinylating a surface can be obtained, for example, from Vector laboratories. Alternatively, biotinylation can be performed with BLCPA: EZ-Link Biotin LC-PEO-Amine (Pierce, Cat. 21347).

25 In some methods, labeled streptavidin (e.g., with a fluorescent label) of very low concentration (e.g., in the  $\mu\text{M}$ , nM or pM range) is used to coat the substrate surface prior to template immobilization. This facilitates immobilization of the template with single molecule resolution. It also allows monitoring of spots on the substrate to which the template molecules are attached, and subsequent nucleotide incorporation events.

30 While diverse polynucleotide templates can be each immobilized to and sequenced in a separate substrate, multiple templates can also be analyzed on a single substrate. In the latter scenario, the templates are attached at different locations on the substrate. This can be accomplished by a variety of different methods, including hybridization of primer capture sequences to oligonucleotides immobilized at different points

on the substrate, and sequential activation of different points down the substrate towards template immobilization.

Methods of creation of surfaces with arrays of oligonucleotides have been described, e.g., in U.S. Patent Nos. 5,744,305, 5,837,832, and 6,077,674. Primers with two domains, a priming domain and a capture domain, can be used to anchor templates to the substrate. The priming domain is complementary to the target template. The capture domain is present on the non-extended side of the priming sequence. It is not complementary to the target template, but rather to a specific oligonucleotide sequence present on the substrate. The target templates can be separately hybridized with their primers, or (if the priming sequences are different) simultaneously hybridized in the same solution. Incubation of the primer/template duplexes with the substrate under hybridization conditions allows attachment of each template to a unique spot. Multiple substrates can be charged with templates in this fashion simultaneously.

Another method for attaching multiple templates to the surface of a single substrate is to sequentially activate portions of the substrate and attach template to them. Activation of the substrate can be achieved by either optical or electrical means. Optical illumination can be used to initiate a photochemical deprotection reaction that allows attachment of the template to the surface (see, e.g., U.S. Patent Nos. 5,599,695, 5,831,070, and 5,959,837). For instance, the substrate surface can be derivitized with "caged biotin", a commercially available derivative of biotin that becomes capable of binding to avidin only after being exposed to light. Templates can then be attached by exposure of a site to light, filling the channel with avidin solution, washing, and then flowing biotinylated template into the channel. Another variation is to prepare avidinylated substrate and a template with a primer with a caged biotin moiety; the template can then be immobilized by flowing into the channel and illumination of the solution above a desired area. Activated template/primer duplexes are then attached to the first wall they diffused to, yielding a diffusion limited spot.

Electrical means can also be used to direct template to specific locations on a substrate. By positively charging one electrode in the channel and negatively charging the others, a field gradient can be created which drives the template to a single electrode, where it can attach (see, e.g., U.S. Patent Nos. 5,632,957, 6,051,380, and 6,071,394). Alternatively, it can be achieved by electrochemically activating regions of the surface and changing the voltage applied to the electrodes. Patterning of particular chemicals, include proteins and DNA is possible with a stamp method, in which a microfabricated plastic stamp is pressed on the surface (see, e.g., Lopez et al., J. Amer. Chem. Soc. 115:10774-81, 1993). Different

templates can also be attached to the surface randomly as the reading of each individual is independent from the others.

E. Treatment of substrate surface

5 In some applications, surface of the substrate is pretreated to create surface chemistry that facilitates attachment of the polynucleotide templates and subsequent synthesis reactions. The surface chemistry also reduces the background from non specific attachment of free labeled nucleotide to the surface of the substrate.

In some methods, the surface is coated with a polyelectrolyte multilayer (PEM). In some methods, non-PEM based surface chemistry can be created prior to template attachment. Preferably, the substrate surface is coated with a polyelectrolyte multilayer (PEM). Attachment of templates to PEM-coated surface can be accomplished by light-directed spatial attachment (see, e.g., U.S. Patent Nos. 5,599,695, 5,831,070, and 5,959,837). Alternatively, the templates can be attached to PEM-coated surface entire chemically (see below for detail).

PEM formation has been described in Decher et al. (Thin Solid Films, 210:831-835, 1992). PEM formation proceeds by the sequential addition of polycations and polyanions, which are polymers with many positive or negative charges, respectively. Upon addition of a polycation to a negatively-charged surface, the polycation deposits on the surface, forming a thin polymer layer and reversing the surface charge. Similarly, a polyanion deposited on a positively charged surface forms a thin layer of polymer and leaves a negatively charged surface. Alternating exposure to poly(+) and poly(-) generates a polyelectrolyte multilayer structure with a surface charge determined by the last polyelectrolyte added; in the case of incompletely-charged surfaces, multiple-layer deposition also tends to increase surface charge to a well defined and stable level.

An exemplified scheme of coating a substrate with PEM for immobilizing polynucleotide is provided in PCT publication WO 01/32930. Detailed procedures are also disclosed in the Examples below. Briefly, the surface of the substrate (e.g., a glass cover slip) is cleaned with a RCA solution. After cleaning, the substrate is coated with a polyelectrolyte multilayer (PEM). Following biotinylation of the carboxylic acid groups, streptavidin is then applied to generate a surface capable of capturing biotinylated molecules. Biotinylated polynucleotide templates are then added to the coated glass cover slip for attachment. The surface chemistry thus created provides various advantages for the methods of the present invention, because it generates a strong negatively-charged surface which

repels the negatively-charged nucleotides. First, a polyelectrolyte multilayer terminated with carboxylic acid-bearing polymer is easy to attach polynucleotide to because carboxylic acids are good targets for covalent bond formation. In addition, the attached template is active for extension by polymerases – most probably, the repulsion of like charges prevents the  
5 template from “laying down” on the surface. Finally, the negative charge repels the fluorescent nucleotides, and nonspecific binding is low.

The attachment scheme described here is easy to generalize on. Without modification, the PEM/biotin/streptavidin surface that is produced can be used to capture or immobilize any biotinylated molecule. A slight modification can be the use of another  
10 capture pair, e.g., substituting digoxigenin (dig) for biotin and labeling the molecule to be immobilized with anti-digoxigenin (anti-dig). Reagents for biotinylation or dig-labeling of amines are all commercially available.

Another generalization is that the chemistry is nearly independent of the surface chemistry of the support. Glass, for instance, can support PEMs terminated with  
15 either positive or negative polymer, and a wide variety of chemistry for either. But other substrates such as silicone, polystyrene, polycarbonate, etc, which are not as strongly charged as glass, can still support PEMs. The charge of the final layer of PEMs on weakly-charged surfaces becomes as high as that of PEMs on strongly-charged surfaces, as long as the PEM has sufficiently-many layers. This means that all the advantages of the  
20 glass/PEM/biotin/Streptavidin/biotin-DNA surface chemistry can be applied to other substrates.

#### IV. Primer Extension Reaction

Once templates are immobilized to the surface of a substrate, primer extension  
25 reactions are performed, e.g., as described in Sambrook, *supra*; Ausubel, *supra*; and Hyman, Anal. Biochem., 174, p. 423, 1988. In some methods, the primer is extended by a polynucleotide polymerase in the presence of a single type of labeled nucleotide. In other methods, all four types of differently labeled nucleotides are present. In some applications of the present invention, a combination of labeled and non-labeled nucleotides are used in the  
30 analysis. A label is incorporated into the template/primer complex only if the specific labeled nucleotide added to the reaction is complementary to the nucleotide on the template adjacent the 3' end of the primer. Optionally, the template is subsequently washed to remove any unincorporated label, and the presence of any incorporated label is determined. As some

errors can be caused by the polymerase, the reaction conditions and incubation time should minimize these errors.

A. Labeled nucleotides

To facilitate detection of nucleotide incorporation, at least one and usually all types of the deoxyribonucleotides (dATP, dTTP, dGTP, dCTP, dUTP/dTTP) or nucleotides (ATP, UTP, GTP, and CTP) are labeled with fluorophores. When more than one type of nucleotides are labeled, a different kind of label can be used to label each different type of nucleotide. However, in some applications, the different types of nucleotides can be labeled with the same kind of labels.

Various fluorescent labels can be used to label the nucleotides in the present invention. The fluorescent label can be selected from any of a number of different moieties. The preferred moiety is a fluorescent group for which detection is quite sensitive. The affinity to the surface could be changed between different dyes. Low affinity to the surface is preferred. For example, Cy3 and Cy5 are used to label the primer or nucleotides in some methods of the invention. However, Cy5 has higher affinity to the surface under certain experimental condition than Cy3.

Other factors that need to be considered include stability of the dyes. For example, Cy5 is less stable and tends to bleach faster than Cy3. Such property can be of advantage or disadvantage, depending on the circumstances. In addition, different sizes of the dyes can also affect efficiency of incorporation of labeled nucleotides. Further, length of the linker between the dye and the nucleotide can impact efficiency of the incorporation (see, Zhu and Waggoner, Cytometry 28: 206, 1997).

An exemplary list of fluorophores, with their corresponding absorption/emission wavelength indicated in parenthesis, that can be used in the present invention include Cy3 (550/565), Cy5 (650/664), Cy7 (750/770), Rho123 (507/529), R6G (528/551), BODIPY 576/589 (576/589), BODIPY TR (588/616), Nile Blue (627/660), BODIPY 650/665 (650/665), Sulfo-IRD700 (680/705), NN382 (778/806), Alexa488 (490/520), Tetramethylrhodamine (550/570). and Rodamine X (575/605).

The fluorescently labeled nucleotides can be obtained commercially (e.g., from NEN DuPont, Amersham, or BDL). Alternatively, fluorescently labeled nucleotides can also be produced by various fluorescence-labeling techniques, e.g., as described in Kambara et al. (1988) "Optimization of Parameters in a DNA Sequenator Using Fluorescence Detection," Bio/Technol. 6:816-821; Smith et al. (1985) Nucl. Acids Res, 13:2399-2412; and

Smith et al. (1986) Nature 321:674-679. Acyl fluoride of Cy5 cyanine dye can also be synthesized and labeled as described in U.S. Patent No. 6,342,326.

There is a great deal of practical guidance available in the literature for providing an exhaustive list of fluorescent and chromogenic molecules and their relevant optical properties (*see, for example, Berlman, Handbook of Fluorescence Spectra of Aromatic Molecules*, 2nd Edition (Academic Press, New York, 1971); Griffiths, *Colour and Constitution of Organic Molecules* (Academic Press, New York, 1976); Bishop, Ed., *Indicators* (Pergamon Press, Oxford, 1972); Haugland, *Handbook of Fluorescent Probes and Research Chemicals* (Molecular Probes, Eugene, 1992) Pringsheim, *Fluorescence and Phosphorescence* (Interscience Publishers, New York, 1949); and the like. Further, there is extensive guidance in the literature for derivatizing fluorophore and quencher molecules for covalent attachment via common reactive groups that can be added to a nucleotide, as exemplified by the following references: Haugland (*supra*); Ullman *et al.*, U.S. Pat. No. 3,996,345; Khanna *et al.*, U.S. Pat. No. 4,351,760.

There are many linking moieties and methodologies for attaching fluorophore moieties to nucleotides, as exemplified by the following references: Eckstein, editor, *Oligonucleotides and Analogues: A Practical Approach* (IRL Press, Oxford, 1991); Zuckerman *et al.*, *Nucleic Acids Research*, 15: 5305-5321 (1987) (3' thiol group on oligonucleotide); Sharma *et al.*, *Nucleic Acids Research*, 19: 3019 (1991) (3' sulfhydryl); Giusti *et al.*, *PCR Methods and Applications*, 2: 223-227 (1993) and Fung *et al.*, U.S. Pat. No. 4,757,141 (5' phosphoamino group via Aminolink<sup>TM</sup>. II available from Applied Biosystems, Foster City, Calif.) Stabinsky, U.S. Pat. No. 4,739,044 (3' aminoalkylphosphoryl group); Agrawal *et al.*, *Tetrahedron Letters*, 31: 1543-1546 (1990) (attachment via phosphoramidate linkages); Sproat *et al.*, *Nucleic Acids Research*, 15: 4837 (1987) (5' mercapto group); Nelson *et al.*, *Nucleic Acids Research*, 17: 7187-7194 (1989) (3' amino group); and the like.

In instances where a multi-labeling scheme is utilized, a wavelength which approximates the mean of the various candidate labels' absorption maxima may be used. Alternatively, multiple excitations may be performed, each using a wavelength corresponding to the absorption maximum of a specific label.

## B. Other reaction reagents

### 1. Polymerases

Many polymerases can be selected for use in this invention. Preferred polymerases are able to tolerate labels on the nucleobase. For example, some applications of the present invention employ polymerases that have increased ability to incorporate modified, fluorophore-labeled, nucleotides into polynucleotides. Examples of such polymerases, e.g., mutant bacteriophage T4 DNA polymerases, have been described in U.S. Patent No. 5,945,312.

Depending on the template, either RNA polymerase, DNA polymerases or reverse transcriptase can be used in the primer extension. For analysis of DNA templates, many DNA polymerases are available. Examples of suitable DNA polymerases include, but are not limited to, Sequenase 2.0.RTM., T4 DNA polymerase or the Klenow fragment of DNA polymerase 1, or Vent polymerase. In some methods, polymerases which lack 3' → 5' exonuclease activity can be used (e.g., T7 DNA polymerase (Amersham) or Klenow -exo fragment of DNA polymerase I (New England Biolabs)). In some methods, when it is desired that the polymerase have proof-reading activity, polymerases lacking 3' → 5' exonuclease activity are not used. In some methods, thermostable polymerases such as ThermoSequenase™ (Amersham) or Taquenase™ (ScienTech, St Louis, MO) are used.

In general, the polymerase should have a fidelity (incorporation accuracy) of at least 99% and a processivity (number of nucleotides incorporated before the enzyme dissociates from the DNA) of at least 20 nucleotides, with greater processivity preferred.

Examples include T7 DNA polymerase, T5 DNA polymerase, HIV reverse transcriptase, *E. coli* DNA pol I, T4 DNA polymerase, T7 RNA polymerase, Taq DNA polymerase and *E. coli* RNA polymerase, Phi29 DNA polymerase.

The nucleotides used in the methods should be compatible with the selected polymerase. Procedures for selecting suitable nucleotide and polymerase combinations can be adapted from Ruth et al. (1981) *Molecular Pharmacology* 20:415-422; Kutateladze, T., et al. (1984) *Nuc. Acids Res.*, 12:1671-1686; Chidgeavadze, Z., et al. (1985) *FEBS Letters*, 183:275-278.



The polymerase can be stored in a separate reservoir and flowed onto the substrates (or into a flow chamber/cell which houses the substrate) prior to each extension reaction cycle. The enzyme can also be stored together with the other reaction agents (e.g., the nucleotide triphosphates). Alternatively, the polymerase can be immobilized onto the surface of the substrate while the polynucleotide template is added to the solution.

## 2. Blocking agents

In some methods, it may be desirable to employ a chain elongation inhibitor in the primer extension reaction (see, e.g., Dower et al., U.S. Patent No. 5,902,723). Chain elongation inhibitors are nucleotide analogues which either are chain terminators which prevent further addition by the polymerase of nucleotides to the 3' end of the chain by becoming incorporated into the chain themselves. In some methods, the chain elongation inhibitors are dideoxynucleotides. Where the chain elongation inhibitors are incorporated into the growing polynucleotide chain, they should be removed after incorporation of the labeled nucleotide has been detected, in order to allow the sequencing reaction to proceed using different labeled nucleotides. Some 3' to 5' exonucleases, e.g., exonuclease III, are able to remove dideoxynucleotides.

Other than chain elongation inhibitors, a blocking agent or blocking group can be employed on the 3' moiety of the deoxyribose group of the labeled nucleotide to prevent nonspecific incorporation. Optimally, the blocking agent should be removable under mild conditions (e.g., photosensitive, weak acid labile, or weak base labile groups), thereby allowing for further elongation of the primer strand with a next synthetic cycle. If the blocking agent also contains the fluorescent label, the dual blocking and labeling functions are achieved without the need for separate reactions for the separate moieties. For example, the labeled nucleotide can be labeled by attachment of a fluorescent dye group to the 3' moiety of the deoxyribose group, and the label is removed by cleaving the fluorescent dye from the nucleotide to generate a 3' hydroxyl group. The fluorescent dye is preferably linked to the deoxyribose by a linker arm which is easily cleaved by chemical or enzymatic means.

Examples of blocking agents include, among others, light sensitive groups such as 6-nitroveratryloxycarbonyl (NVOC), 2-nitrobenzyloxycarbonyl (NBOC),  $\alpha,\alpha$ -dimethyl-dimethoxybenzyloxycarbonyl (DDZ), 5-bromo-7-nitroindolinyl, o-hydroxy-2-methyl cinnamoyl, 2-oxymethylene anthraquinone, and t-butyl oxycarbonyl (TBOC). Other blocking reagents are discussed, e.g., in U.S. Ser. No. 07/492,462; Patchornik (1970) J. Amer. Chem. Soc. 92:6333; and Amit et al. (1974) J. Org. Chem. 39:192. Nucleotides

possessing various labels and blocking groups can be readily synthesized. Labeling moieties are attached at appropriate sites on the nucleotide using chemistry and conditions as described, e.g., in Gait (1984) Oligonucleotide Synthesis: A Practical Approach, IRL Press, Oxford.

5 C. Reaction conditions

The reaction mixture for the sequencing comprises an aqueous buffer medium which is optimized for the particular polymerase. In general, the buffer includes a source of monovalent ions, a source of divalent cations and a buffering agent. Any convenient source of monovalent ions, such as KCl, K-acetate, NH<sub>4</sub>-acetate, K-glutamate, NH<sub>4</sub>Cl, ammonium sulfate, and the like may be employed, where the amount of monovalent ion source present in  
10 the buffer will typically be present in an amount sufficient to provide for a conductivity in a range from about 500 to 20,000, usually from about 1000 to 10,000, and more usually from about 3,000 to 6,000 micromhos.

The divalent cation may be magnesium, manganese, zinc and the like, where  
15 the cation will typically be magnesium. Any convenient source of magnesium cation may be employed, including MgCl<sub>2</sub>, Mg-acetate, and the like. The amount of Mg ion present in the buffer may range from 0.5 to 20 mM, but will preferably range from about 1 to 12mM, more preferably from 2 to 10mM and will ideally be about 5mM.

Representative buffering agents or salts that may be present in the buffer  
20 include Tris, Tricine, HEPES, MOPS and the like, where the amount of buffering agent will typically range from about 5 to 150 mM, usually from about 10 to 100 mM, and more usually from about 20 to 50 mM, where in certain preferred embodiments the buffering agent will be present in an amount sufficient to provide a pH ranging from about 6.0 to 9.5, where most preferred is pH 7.6 at 25° C. Other agents which may be present in the buffer medium include  
25 chelating agents, such as EDTA, EGTA and the like.

D. Removal of labels and blocking group

By repeating the incorporation and label detection steps until incorporation is detected, the nucleotide on the template adjacent the 3' end of the primer can be identified.  
30 Once this has been achieved, the label should be removed before repeating the process to discover the identity of the next nucleotide. Removal of the label can be effected by removal of the labeled nucleotide using a 3'-5' exonuclease and subsequent replacement with an unlabeled nucleotide. Alternatively, the labeling group can be removed from the nucleotide.

Release of the fluorescence dye can be achieved if a detachable connection between the nucleotide and the fluorescence molecule is used. For example, the use of disulfide bonds enables one to disconnect the dye by applying a reducing agent like dithiothreitol (DTT). In a further alternative, where the label is a fluorescent label, it is possible to neutralize the label by bleaching it with radiation. Photobleaching can be performed according to methods, e.g., as described in Jacobson et al., "International Workshop on the Application of Fluorescence Photobleaching Techniques to Problems in Cell Biology", Federation Proceedings, 42:72-79, 1973; Okabe et al., J Cell Biol 120:1177-86, 1993; Wedekind et al., J Microsc. 176 Pt 1): 23-33, 1994; and Close et al., Radiat Res 53:349-57, 1973.

If chain terminators or 3' blocking groups have been used, these should be removed before the next cycle can take place. 3' blocking groups can be removed by chemical or enzymatic cleavage of the blocking group from the nucleotide. For example, chain terminators are removed with a 3'-5' exonuclease, e.g., exonuclease III. Once the label and terminators/blocking groups have been removed, the cycle is repeated to discover the identity of the next nucleotide.

#### E. Sample housing.

The solid substrate is optionally housed in a flow chamber having an inlet and outlet to allow for renewal of reactants which flow past the immobilized moieties. The flow chamber can be made of plastic or glass and should either be open or transparent in the plane viewed by the microscope or optical reader. Electro-osmotic flow requires a fixed charge on the solid substrate and a voltage gradient (current) passing between two electrodes placed at opposing ends of the solid support. Pressure driven flow can be facilitated by microfluidic device with an external pressure source or by microfluidic peristaltic pump (see, e.g., Unger et al., Science 288: 113-116, 2000).

The flow chamber can be divided into multiple channels for separate sequencing. Examples of micro flow chambers are described in Fu *et al.* (*Nat. Biotechnol.* (1999) 17:1109) which describe a microfabricated fluorescence-activated cell sorter with 3µm x 4µm channels that utilizes electro-osmotic flow for sorting. Preferably, the flow chamber contains microfabricated synthesis channels as described in WO01/32930. The polynucleotide templates can be immobilized to the surface of the synthesis channels. These synthesis channels can be in fluid communication with a microfluidic device which controls flow of reaction reagents. Preferred microfluidic devices that can be employed to control flow of reaction reagents in the present invention have been described in WO01/32930.

The present invention also provide apparatus for carrying out the methods of the invention. Other than the substrate to which the target polynucleotides or primers are attached, the apparatus usually comprise a flow chamber in which the substrate is housed. In addition, the apparatus can optionally contain plumbing devices (e.g., an inlet and an outlet port), a light source, and a detection system described herein. Preferably, a microfabricated apparatus as described in WO01/32930 is adapted to house the substrate of the present invention.

V. Detection of Incorporated Signals

A. Detection system in general

Methods for visualizing single molecules of DNA labeled with an intercalating dye include, e.g., fluorescence microscopy as described in Houseal et al., *Biophysical Journal* 56: 507, 1989. While usually signals from a plurality of molecules are to be detected with the sequencing methods of the present invention, fluorescence from single fluorescent dye molecules can also be detected. For example, a number of methods are available for this purpose (see, e.g., Nie et al., *Science* 266: 1013, 1994; Funatsu et al., *Nature* 374: 555, 1995; Mertz et al., *Optics Letters* 20: 2532, 1995; and Unger et al., *Biotechniques* 27:1008, 1999). Even the fluorescent spectrum and lifetime of a single molecule excited-state can be measured (Macklin et al., *Science* 272: 255, 1996). Standard detectors such as a photomultiplier tube or avalanche photodiode can be used. Full field imaging with a two stage image intensified CCD camera can also used (Funatsu et al., supra). Low noise cooled CCD can also be used to detect single fluorescence molecules (see, e.g., Unger et al., *Biotechniques* 27: 1008-1013, 1999; and SenSys spec: <http://www.photomet.com/pdfs/datasheets/sensys/ss1401e.pdf>).

The detection system for the signal or label can also depend upon the label used, which can be defined by the chemistry available. For optical signals, a combination of an optical fiber or charged couple device (CCD) can be used in the detection step. In those circumstances where the matrix is itself transparent to the radiation used, it is possible to have an incident light beam pass through the substrate with the detector located opposite the substrate from the polynucleotides. For electromagnetic labels, various forms of spectroscopy systems can be used. Various physical orientations for the detection system are available and discussion of important design parameters is provided in the art (e.g., Arndt-Jovin et al., *J Cell Biol* 101: 1422-33, 1985; and Marriott et al., *Biophys J* 60: 1374-87, 1991).

Many applications of the invention require the detection of incorporation of fluorescently labeled nucleotides into single template molecules in a solution. The single-molecule fluorescence detection of the present invention can be practiced using optical setups including near-field scanning microscopy, far-field confocal microscopy, wide-field epi-illumination, and total internal reflection fluorescence (TIRF) microscopy. General reviews are available describing this technology, including, e.g., Basche et. al., eds., 1996, Single molecule optical detection, imaging, and spectroscopy, Weinheim:VCM; and Plakhotnik, et. al., Single-molecule spectroscopy, *Ann. Rev. Phys. Chem.* 48: 181-212. In general, the methods involve detection of laser activated fluorescence using microscope equipped with a camera. It is sometimes referred to as a high-efficiency photon detection system (see, e.g., Nie, et. al., 1994, Probing individual molecules with confocal fluorescence microscopy, *Science* 266:1018-1019. Other suitable detection systems are discussed in the Examples below.

Suitable photon detection systems include, but are not limited to, photodiodes and intensified CCD cameras. In a preferred embodiment, an intensified charge couple device (ICCD) camera is used. The use of a ICCD camera to image individual fluorescent dye molecules in a fluid near the surface of the glass slide is advantageous for several reasons. With an ICCD optical setup, it is possible to acquire a sequence of images (movies) of fluorophores. In certain aspects, each of the dNTPs or NTPs employed in the methods has a unique fluorophore associated with it, as such, a four-color instrument can be used having four cameras and four excitation lasers. Preferably the image could be split to four quarters and imaged by a single camera. For example, the micro-imager of Optical Insights LTD is a simple device that splits the image to four different images in four different spectra just in front of the port of the camera. Illumination with only one laser excitation for the four colors is possible if suitable dyes are used (see, e.g., Rosenblum et al, *Nucleic Acids Research* 25:4500, 1997). For example, the BigDyes have single excitation wavelength spectrum and four different emission wavelength spectrums. They can be obtained from Applied Biosystems (see, <http://www.appliedbiosystems.com/products/productdetail.cfm?ID=82>). Nanocrystals are also found to have a variety of emission wavelengths for a given excitation (see, e.g., U.S. Patent No. 6,309,701; and Lacoste et al., *Proc. Natl. Acad. Sci. USA* 97: 9461-6, 2000). Thus, it is possible to use such optical setup to sequence DNA. In addition, many different DNA molecules spread on a solid support (e.g., a microscope slide) can be imaged and sequenced simultaneously.

B. Total internal reflection fluorescence (TIRF) microscopy

In some preferred embodiments, the present invention uses total internal reflection fluorescence (TIRF) microscopy for two-dimensional imaging fluorescence detection. TIRF microscopy is well known in the art. See, e.g., Watkins et al., J Biomed Mater Res 11:915-38, 1977; and Axelrod et al., J Microsc, 129:19-28, 1983. TIRF microscopy uses totally internally reflected excitation light. When a laser beam was totally reflected at the interface between a liquid and a solid substrate (e.g., a glass), the excitation light beam penetrates only a short distance into the liquid. In other words, the optical field does not end abruptly at the reflective interface, but its intensity falls off exponentially with distance. This surface electromagnetic field, called the 'evanescent wave', can selectively excite fluorescent molecules in the liquid near the interface. The thin evanescent optical field at the interface provides low background and enables the detection of single molecules with high signal-to-noise ratio at visible wavelengths (see, M. Tokunaga et al., *Biochem. and Biophys. Res. Comm.* 235, 47 (1997) and P. Ambrose, *Cytometry*, 36, 244 (1999)).

TIRF microscopy has been used to examine various molecular or cellular activities, e.g., cell/substrate contact regions of primary cultured rat myotubes with acetylcholine receptors labeled by fluorescent alpha-bungarotoxin, and human skin fibroblasts labeled with a membrane-incorporated fluorescent lipid (see, e.g., Thompson et al., *Biophys J.* 33:435-54, 1981; Axelrod, *J. Cell. Biol.* 89: 141-5, 1981; and Burghardt et al., *Biochemistry* 22:979-85, 1983). TIRF examination of cell/surface contacts dramatically reduces background from surface autofluorescence and debris. TIRF has also been combined with fluorescence photobleaching recovery and correlation spectroscopy to measure the chemical kinetic binding rates and surface diffusion constant of fluorescent labeled serum protein binding (at equilibrium) to a surface (see, e.g., Burghardt et al., *Biophys J.* 33:455-67, 1981); and Thompson et al., *Biophys J.* 43:103-14, 1983). Additional examples of TIRF detection of single molecules have been described in Vale et. al., 1996, Direct observation of single kinesin molecules moving along microtubules, *Nature* 380: 451; and Xu et al., 1997, Direct Measurement of Single-Molecule Diffusion and Photodecomposition in Free Solution, *Science* 275: 1106-1109.

The penetration of the field beyond the glass depends on the wavelength and the laser beam angle of incidence. Deeper penetrance is obtained for longer wavelengths and for smaller angles to the surface normal within the limit of a critical angle. In typical assays, fluorophores are detected within about 200 nm from the surface which corresponds to the contour length of about 600 base pairs of DNA. In some embodiments, when longer

polynucleotide templates are analyzed, the polymerase rather than the template is immobilized to the surface so the reaction occurs near the surface at all time. In some embodiments, a prism-type TIRF geometry for single-molecule imaging as described by Xu and Yeung is used (*see*, X-H.N. Xu *et al.*, *Science*, **281**, 1650 (1998)). In some embodiments, an objective type TIRF is used to provide space above the objective so that a microfluidic device can be used (*see*, e.g., Tokunaga *et al.*, *Biochem Biophys Res Commun* 235: 47-53, 1997; Ambrose *et al.*, *Cytometry* 36:224;1999; and Braslavsky *et al.*, *Applied Optics* 40:5650, 2001).

Total internal reflection can be utilized with high numerical aperture objectives (ranging between 1.4 and 1.65 in aperture), preferentially using an inverted microscope. The numerical aperture of an objective is a function of the max angle that can be collected (or illuminated) with the objective in a given refractive index of the media (i.e.,  $NA = n \cdot \sin(\theta_{\text{Max}})$ ). If  $\theta_{\text{Max}}$  is larger than  $\theta_{\text{Critic}}$  for reflection, some of the illuminated rays will be totally internal reflected. So using the peripheral of a large NA objective one can illuminate the sample with TIR through the objective and use the same objective to collect the fluorescence light. Therefore, the objective plays double roles as a condenser and an imaging objective.

Single molecule detection can be achieved using flow cytometry where flowing samples are passed through a focused laser with a spatial filter used to define a small volume. US Pat. No. 4,979,824 describes a device for this purpose. US Pat. No. 4,793,705 describes a detection system for identifying individual molecules in a flow train of the particles in a flow cell. It further describes methods of arranging a plurality of lasers, filters and detectors for detecting different fluorescent nucleic acid base-specific labels. US Pat. No. 4,962,037 also describes a method for detecting an ordered train of labeled nucleotides for obtaining DNA and RNA sequences using an exonuclease to cleave the bases. Single molecule detection on solid supports is also described in Ishikawa, *et al.* (1994) Single-molecule detection by laser-induced fluorescence technique with a position-sensitive photon-counting apparatus, *Jan. J. Appl. Phys.* 33:1571-1576. Ishikawa describes a typical apparatus involving a photon-counting camera system attached to a fluorescence microscope. Lee *et al.* (*Anal. Chem.*, 66:4142-4149, 1994) describes an apparatus for detecting single molecules in a quartz capillary tube. The selection of lasers is dependent on the label and the quality of light required. Diode, helium neon, argon ion, argon-krypton mixed ion, and double Nd:YAG lasers are useful in this invention.

C. Excitation and scanning

In some applications, fluorescent excitation is exerted with a Q-switched frequency doubled Nd YAG laser, which has a KHz repetition rate, allowing many samples to be taken per second. For example, a wavelength of 532 nm is ideal for the excitation of rhodamine. It is a standard device that has been used in the single molecule detection scheme (Smith et al., *Science* 253:1122, 1992). A pulsed laser allows time resolved experiments, which are useful for rejecting extraneous noise. In some methods, excitation can be performed with a mercury lamp and signals from the incorporated nucleotides can be detected with an CCD camera (see, e.g., Unger et al., *Biotechniques* 27:1008, 1999).

Incorporated signals can be detected by scanning the substrates. The substrates can be scanned simultaneously or serially, depending on the scanning method used. The signals can be scanned using a CCD camera (TE/CCD512SF, Princeton Instruments, Trenton, N.J.) with suitable optics (Ploem, J. S., in *Fluorescent and Luminescent Probes for Biological Activity*, Mason, T. W., Ed., Academic Press, London, pp. 1-11, 1993), such as described in Yershov et al. (*Proc. Natl. Acad. Sci.* 93:4913, 1996), or can be imaged by TV monitoring (Khrapko et al., *DNA Sequencing* 1:375, 1991). The scanning system should be able to reproducibly scan the substrates. Where appropriate, e.g., for a two dimensional substrate where the substrates are localized to positions thereon, the scanning system should positionally define the substrates attached thereon to a reproducible coordinate system. It is important that the positional identification of substrates be repeatable in successive scan steps.

Various scanning systems can be employed in the methods and apparatus of the present invention. For example, electro-optical scanning devices described in, e.g., U.S. Pat. No. 5,143,854, are suitable for use with the present invention. The system could exhibit many of the features of photographic scanners, digitizers or even compact disk reading devices. For example, a model no. PM500-A1 x-y translation table manufactured by Newport Corporation can be attached to a detector unit. The x-y translation table is connected to and controlled by an appropriately programmed digital computer such as an IBM PC/AT or AT compatible computer. The detection system can be a model no. R943-02 photomultiplier tube manufactured by Hamamatsu, attached to a preamplifier, e.g., a model no. SR440 manufactured by Stanford Research Systems, and to a photon counter, e.g., an SR430 manufactured by Stanford Research System, or a multichannel detection device. Although a digital signal can usually be preferred, there can be circumstances where analog signals would be advantageous.



The stability and reproducibility of the positional localization in scanning determine, to a large extent, the resolution for separating closely positioned polynucleotide clusters on a two dimensional substrate. Since the successive monitoring at a given position depends upon the ability to map the results of a reaction cycle to its effect on a positionally mapped polynucleotides, high resolution scanning is preferred. As the resolution increases, the upper limit to the number of possible polynucleotides which can be sequenced on a single matrix also increases. Crude scanning systems can resolve only on the order of 1000  $\mu\text{m}$ , refined scanning systems can resolve on the order of 100  $\mu\text{m}$ , more refined systems can resolve on the order of about 10  $\mu\text{m}$ , and with optical magnification systems a resolution on the order of 1.0  $\mu\text{m}$  is available. The limitations on the resolution can be diffraction limited and advantages can arise from using shorter wavelength radiation for fluorescent scanning steps. However, with increased resolution, the time required to fully scan a matrix can increased and a compromise between speed and resolution can be selected. Parallel detection devices which provide high resolution with shorter scan times are applicable where multiple detectors are moved in parallel.

In some applications, resolution often is not so important and sensitivity is emphasized. However, the reliability of a signal can be pre-selected by counting photons and continuing to count for a longer period at positions where intensity of signal is lower.

Although this decreases scan speed, it can increase reliability of the signal determination.

Various signal detection and processing algorithms can be incorporated into the detection system. In some methods, the distribution of signal intensities of pixels across the region of signal are evaluated to determine whether the distribution of intensities corresponds to a time positive signal.

#### D. Detection of incorporation of multiple fluorescent labels: FRET

In some aspects of the present application, incorporation of different types of nucleotides into a primer is detected using different fluorescent labels on the different types of nucleotides. When two different labels are incorporated into the primer in close vicinity, signals due to fluorescence resonance energy transfer (FRET) can be detected. FRET is a phenomenon that has been well documented in the literature, e.g., in T. Foster, Modern Quantum Chemistry, Istanbul Lectures, Part III, 93-137, 1965, Academic Press, New York; and Selvin, "Fluorescence Resonance Energy Transfer," Methods in Enzymology 246: 300-335, 1995. In FRET, one of the fluorophores (donor) has an emission spectrum that overlaps the excitation spectrum of the other fluorophore (acceptor) and transfer of energy takes place

from the donor to the acceptor through fluorescence resonance energy transfer. The energy transfer is mediated by dipole-dipole interaction. Spectroscopically, when the donor is excited, its specific emission intensity decreases while the acceptor's specific emission intensity increases, resulting in fluorescence enhancement.

5           Detection of single molecule FRET signal reveals sequence information and facilitates interpretation of the sequencing data. Detection of FRET signal in the present invention can be performed accordingly to various methods described in the art (e.g., US Patent No. 5,776,782). FRET has been used to studying various biological activities of biomacromolecules including polynucleotides. For example, Cooper et al. disclosed  
10 fluorescence energy transfer in duplex and branched DNA molecules (Biochemistry 29: 9261-9268, 1990). Lazowski et al. reported highly sensitive detection of hybridization of oligonucleotides to specific sequences of nucleic acids by FRET (Antisense Nucleic Acid Drug Dev. 10: 97-103, 2000). Methods for nucleic acid analysis using FRET were also described in US Patent Nos. 6,177,249 and 5,945,283. Efficacy of using FRET to detect  
15 multiple nucleotides incorporation into single polynucleotide molecules is also exemplified in Example 8 of the present application.

Any of a number of fluorophore combinations can be selected for labeling the nucleotides in the present invention for detection of FRET signals (see for example, Pesce et al., eds, Fluorescence Spectroscopy, Marcel Dekker, New York, 1971; White et al.,  
20 Fluorescence Analysis: A practical Approach, Marcel Dekker, New York, 1970; Handbook of Fluorescent Probes and Research Chemicals, 6th Ed, Molecular Probes, Inc., Eugene, Oreg., 1996; which are incorporated by reference). In general, a preferred donor fluorophore is selected that has a substantial spectrum of the acceptor fluorophore. Furthermore, it may also be desirable in certain applications that the donor have an excitation maximum near a  
25 laser frequency such as Helium-Cadmium 442 nm or Argon 488 nm. In such applications the use of intense laser light can serve as an effective means to excite the donor fluorophore. The acceptor fluorophore has a substantial overlap of its excitation spectrum with the emission spectrum of the donor fluorophore. In addition, the wavelength of the maximum of the emission spectrum of the acceptor moiety is preferably at least 10 nm greater than the  
30 wavelength of the maximum of the excitation spectrum of the donor moiety. The emission spectrum of the acceptor fluorophore is shifted compared to the donor spectrum.

Suitable donors and acceptors operating on the principle of fluorescence energy transfer (FET) include, but are not limited to, 4-acetamido-4'-isothiocyanatostilbene-2,2'-disulfonic acid; acridine and derivatives: acridine, acridine isothiocyanate; 5-(2'-

aminoethyl)aminonaphthalene-1-sulfonic acid (EDANS); 4-amino-N-[3-  
 vinylsulfonyl]phenyl]naphthalimide-3,5 disulfonate; N-(4-anilino-1-naphthyl)maleimide;  
 anthranilamide; BODIPY; Brilliant Yellow; coumarin and derivatives: coumarin, 7-amino-4-  
 methylcoumarin (AMC, Coumarin 120), 7-amino-4-trifluoromethylcoumarin (Coumaran  
 5 151); cyanine dyes; cyanosine; 4',6-diaminidino-2-phenylindole (DAPI); 5', 5''-  
 dibromopyrogallol-sulfonaphthalein (Bromopyrogallol Red); 7-diethylamino-3-(4'-  
 isothiocyanatophenyl)-4-methylcoumarin; diethylenetriamine pentaacetate; 4,4'-  
 diisothiocyanatodihydro-stilbene-2,2'-disulfonic acid; 4,4'-diisothiocyanatostilbene-2,2'-  
 disulfonic acid; 5-[dimethylamino]naphthalene-1-sulfonyl chloride (DNS, dansylchloride); 4-  
 10 dimethylaminophenylazophenyl-4'-isothiocyanate (DABITC); eosin and derivatives: eosin,  
 eosin isothiocyanate, erythrosin and derivatives: erythrosin B, erythrosin, isothiocyanate;  
 ethidium; fluorescein and derivatives: 5-carboxyfluorescein (FAM), 5-(4,6-dichlorotriazin-2-  
 yl)aminofluorescein (DTAF), 2',7'-dimethoxy-4'5'-dichloro-6-carboxyfluorescein (JOE),  
 fluorescein, fluorescein isothiocyanate, QFITC, (XRITC); fluorescamine; IR144; IR1446;  
 15 Malachite Green isothiocyanate; 4-methylumbelliferoneortho cresolphthalein; nitrotyrosine;  
 pararosaniline; Phenol Red; B-phycoerythrin; o-phthaldialdehyde; pyrene and derivatives:  
 pyrene, pyrene butyrate, succinimidyl 1-pyrene; butyrate quantum dots; Reactive Red 4  
 (Cibacron™ Brilliant Red 3B-A) rhodamine and derivatives: 6-carboxy-X-rhodamine  
 (ROX), 6-carboxyrhodamine (R6G), lissamine rhodamine B sulfonyl chloride rhodamine  
 20 (Rhod), rhodamine B, rhodamine 123, rhodamine X isothiocyanate, sulforhodamine B,  
 sulforhodamine 101, sulfonyl chloride derivative of sulforhodamine 101 (Texas Red);  
 N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA); tetramethyl rhodamine; tetramethyl  
 rhodamine isothiocyanate (TRITC); riboflavin; rosolic acid; terbium chelate derivatives; Cy  
 3; Cy 5; Cy 5.5; Cy 7; IRD 700; IRD 800; La Jolla Blue; phthalocyanine; and naphthalo  
 25 cyanine.

\*\*\*

Many modifications and variations of this invention can be made without  
 departing from its spirit and scope. The specific embodiments described below are for  
 illustration only and are not intended to limit the invention in any way.

EXAMPLESExample 1 Basic Materials and Methods

- 5 1. Materials and Reaction Reagents
- (1) Solutions and buffers
- RCA: H<sub>2</sub>O:NH<sub>4</sub>OH:H<sub>2</sub>O<sub>2</sub> (6:4:1) boiling for an hour.
- PEI: PolyEthylenImine (Sigma P-3143) (positive charged)
- PALL: Poly(allylamine hydrochloride) (Sigma 283223)
- 10 PACr: Poly(acrylic acid, sodium salt) (Sigma 416045) (negative charged)
- EDC: 9.6mg/ml; 50mM (x10) 1-{3-(Dimethylamino)propyl}-3-ethylcarbodiimide, hydrochloride), Activator for the BLCPA (Sigma- 161462)
- BLCPA: EZ-Link Biotin LC-PEO-Amine (Pierce 21347)
- Stock solution 50mM in MES 10mM (21mg/ml) (x10)
- 15 Streptavidin plus - 1mg/ml in Tris. PROzyme, Code: SA20 (x10)
- Buffers:
- MES (N-morpholinoethanesulfonic acid) PH 5.5 1M (100x)
- TRIS 10mM
- 20 TRIS-MgCl<sub>2</sub> 10mM Tris, 100mM MgCl<sub>2</sub> (x1)
- TKMC (10mM Tris• HCl, 10mM KCl, 10mM MgCl<sub>2</sub>, 5mM Ca Cl<sub>2</sub>, pH 7.0)
- EcoPol : 10mM Tris• HCl, 5mM MgCl<sub>2</sub>, 7.5 mM DTT pH @ 25<sup>0</sup>C; buffer come with the polymerase at (x10)
- 25 (2) Other materials and reagents
- Nucleotides: dTTP, dGTP, dATP, and dCTP-Cy3 at 10μM concentration
- Polymerase: a) Klenow Polymerase I (5 units/μl), New England BioLabs Cat. 210S
- b) Klenow -exo, New England BioLabs Cat. 212S
- 30 c) TAQ
- d) Sequenase
- Hybridization Chamber: Sigma H-1409
- Polynucleotide templates and primers:
- 7G: Biotin - 5'-tcagtcacga gtcacagtc atcagtcac agtcacagtc catcagtcac
- 35 cagtcacag tcacagtcac tcagtcacga gtcacACAC GGAGGTTCTA - 3' (SEQ ID NO:1)
- Primer p7G: 5'- TAGAACCTCCGTGT - 3' (SEQ ID NO:2); the primer can be labeled with Cy5 or Cy3.
- Mu50: Biotin 5'- ctccagcgtgtttatctctgcgagcataatgcctgcgtcatccgccagc 3' (SEQ
- 40 ID NO:3)
- Cy5 labeled primer (PMu50Cy5): Cy5 5' - gctggcgatgac - 3' (SEQ ID NO:4)
- 7G7A - Biotin-5'-
- tttGcttcttAttctttGcttcttAttctttGcttcttAttctttGcttcttAttctttGcttcttAttctttGcttcttAttcttACACGGA
- 45 GGTCTA - 3' (SEQ ID NO:5)

6TA6CG: Biotin- 5' –

ccAtttttGccccccAtttttGccccccAtttttGccccccAtttttGccccccAtttttA-CACGGAGGTTCTA -  
3', (SEQ ID NO:6)

## 5 2. Substrate treatment and template attachment

A fused silica microscope slide (1 mm thick, 25x75 mm size, Esco Cat. R130110) was used to attach DNA templates. The slides was first cleaned with the RCA method as described above and in WO 01/32930. Multilayer of polyallylamine /polyAcrylic were absorbed to the slide. An EZ link connector was then attached to the slides as follows:  
10 the slide was dried, scratched with diamond pencil, and then covered with a hybridization chamber. 120 µl of a mixture of 1:1:8 EDC: BLCPA: MES (50mM EDC, 50mM BLCPA, 10mM MES) was applied to each slide. Following incubation for 20 minutes, 120 µl of Streptavidin Plus diluted to 0.1mg/ml was added to the slide. After 20 min of incubation, the slide was washed with 200µl of Tris 10mM.

15 Preparation of 10pM Oligo: the 7G oligonucleotide template (SEQ ID NO:1) was pre-hybridized with Cy5-labeled primer (SEQ ID NO:2) (in stock at 7µM) in TRIS-MgCl<sub>2</sub> buffer. The treated slide was examined for contamination with the TIR microscope. 200µl of the oligonucleotide/primer mixture was applied to each slide. Following incubation for 10 min, the slide was washed with 200µl ml of Tris 10mM.

20 Addition of nucleotides and polymerase: nucleotides dTTP, dATP, dGTP, and Cy3-dCTP each of 20-100nM were mixed in the ECOPOL buffer. 1 µl Klenow 210S from stock solution (kept in -20°C) was added to 200 microliters of the nucleotide mixture. 120µl of the mixture was then added on each slide. After incubation for 0 to 30 min (for different experiments), the slide was examined with the TIR microscope. Unless otherwise noted, all  
25 reactions were performed at room temperature, while the reaction reagents were kept at 4°C or -20°C. The primer/oligonucleotide hybridization reaction was carried out with a thermocycler machine.

Single molecule resolution was achieve by using very low concentration of the polynucleotide template which ensured that only one template molecule is attached to a  
30 distinct spot on the slide. Single molecule attachment to a distinct is also confirmed by the observation of single bleaching pattern of the attached fluorophores. In the reaction described above, a concentration of about 10pM of a 80-mer oligonucleotide template was

used for immobilizing to the slide. The space between different DNA molecules attached to the surface slide was measured at a few micrometers.

### 3. Imagine with single molecule resolution

As illustrate in Figure 1, the single stranded oligonucleotide template (SEQ ID NO:1) primed with a Cy5 labeled primer sequence (SEQ ID NO:2) was immobilized at a single molecule resolution to the surface of a silica slide using a biotin-streptavidin bond. The surface is coated with polymers on which biotin (EZ link) is tethered. The oligonucleotide template, with a biotin molecule attached to one of its ends, was able to attach to the streptavidin-linked surface. The slide surface was negatively charged which helps to repeal unbound nucleotides. The DNA is specifically attached to the surface by its 5' side, meaning that the primer -which the polymerase extends- is away from the surface.

The template and incorporation of labeled nucleotides were visualized by fluorescence imaging. Location of the oligonucleotide was monitored by fluorescence from the Cy5 labeled primer (SEQ ID NO:2). Incorporation of nucleotides was detected because the nucleotides were labeled with Cy3. After incorporation, the incorporated labels were illuminated. Illumination of Cy3 was at a wavelength of 532nm. Following a typical time of a few seconds of continued illumination, the signals were bleached, typically in a single step.

As shown in Figure 2, imaging of fluorescent signals with single molecule resolution was enabled with surface illumination by total internal reflection (TIR). Ishijima et al. (Cell 92:161-71, 1998) showed that it is possible to observe the fluorescence of single molecules immobilized to a surface in a wet environment even when there are free molecules in the solution. Here, the TIR was facilitated by a dove prism coupling of the laser beam to the silica slide surface. An upright microscope with an immersion oil objective was used to image the surface with an intensified CCD (PentaMax). A filter set (Chroma) was used to reject the illumination frequency and let the fluorescence frequency to reach the ICCD.

#### Example 2 Test for Specific Attachment of Template Molecules to Substrate Surface

This experiment was performed to determine whether the polynucleotide templates are attached to the surface as desired. Figure 3 shows that streptavidin is required for binding the template to the surface and hence detection of incorporated fluorescence signal. The left panel shows that there is no fluorescence signal when only streptavidin-attached surface but no fluorescent labels were present. The middle panel shows that there is no incorporated fluorescent signals when no streptavidin was present on the surface to attach

biotin-labeled oligonucleotide template, even though Cy5-labeled primer was present. The right panel shows that detection of incorporated fluorescent signal when the streptavidin-attached surface, labeled primers, and biotin-labeled oligonucleotide template were present.

5    Example 3.    Determining Processivity of DNA Polymerase in the Presence of Labeled Nucleotides

To determine whether the DNA polymerase accurately incorporates labeled nucleotides into the template, a bulk extension experiment was performed in a test tube rather than on the surface of a substrate. As shown in Figure 5, the results indicate that the  
10    polymerase incorporate all the labeled nucleotides into the correct positions. In this experiment, incorporation of dCTP-Cy3 and a polymerization terminator, ddCTP, were detected using a 7G DNA template (a DNA strand having a G residue every 7 bases; SEQ ID NO:1). The annealed primer was extended in the presence of non-labeled dATP, dGTP, dTTP, Cy3-labeled dCTP, and ddCTP. The ratio of Cy3-dCTP and ddCTP was 3:1. The  
15    reaction products were separated on a gel, fluorescence excited, and the signals detected, using an automatic sequencer ABI-377. The results reveal that incorporation of Cy3-dCTP did not interfere with further extension of the primer along the 7G oligomer template.

Figure 5 shows fluorescence intensity from primer extension products of various lengths which were terminated by incorporation of ddCTP at the different G residues  
20    in the 7G oligomer template (SEQ ID NO:1). The first band is the end of the gel and should not be counted as it is in the very beginning of the gel. The full length of the template is 100 residues. The first band (marked "1" in the graph) corresponds to extension products which were terminated by incorporation of non-labeled ddCTP at the second G residue (position 27) and has incorporated Cy3-dCTP at the first G residue (position 20). Similarly, the tenth band  
25    (marked "10" in the graph) represents extension products which were terminated by incorporation of non-labeled ddCTP at the 10th G residue (position 90) and has incorporated Cy3-dCTP at the previous G residue (i.e., positions 20, 27, 34, 41, 48, 55, 62, 69, 76, and 83). The results showed a nice agreement between the expected positions for Cy3 incorporation in the polynucleotide template and the positions of the fluorescence intensity bands.

30

Example 4.    Detection of single nucleotide incorporation by TIR

Total internal reflection (TIR) fluorescence microscopy allows detection of real-time incorporation of labeled nucleotide into single immobilized polynucleotide template. This illumination method reduce the background from the sample by illuminating

only a thin layer (e.g., in the order of 150nm) near the surface. Even in the presence of free dyes in the solution (up to 50nM), single molecules can be observed. Using TIR, we visualized single molecules of labeled nucleotide bound to DNA in the presence of up to 50nM free dye in solution. Though this concentration is low compared to the concentration  
5 needed for a high rate of incorporation of nucleotides by the DNA polymerase, it was sufficient for its operation.

### 1. Optical setup

The lasers source is shown in Figure 2, the light sources (e.g., laser) are  
10 coupled to the surface by prism. The surface is imaged by a regular 1.3NA microscope objective onto an Intensified CCD (Pentamax). A fluorescent filter in the optical way block the laser intensity and allow the fluorescent signals from the dye molecules pass through(Chroma filters). Optionally, the camera and the shutters for the lasers are controlled  
15 by the computer.

### 2. Illumination

As shown in Figure 6, TIR illumination of polynucleotide-attached slide produced a low background and allowed detection of signals only from immobilized labels. The refraction index of the fused silica glass and the oil beneath the surface is about 1.46.  
20 The refraction index of the liquid above the glass is about 1.33 to 1.35. At the interface of the glass and the water the illumination ray was refracted. If the illumination is very shallow, 70-75 degree from the surface orthogonal, the refracted light was reflected back and not continued in the liquid phase as the critical angel for total internal reflection is about 65-67 degrees ( $\text{TetaCritical} = \sin^{-1}(n_1/n_2)$ ).

25 The illumination process, called evanescent illumination, leaves a decay field near the interface which illuminates only about 150 nm into the liquid phase. Fluorophores dyes can be excited by this field. So only the dyes which are near the surface will emit. Furthermore, free labeled nucleotide molecules in the solution will move around due to Brownian motion. The fast movement of these free molecules produces only a smear signal  
30 because the integration time is in the order of hundred millisecond. Thus, the total internal reflection illumination leads to a low back ground from the free molecules, and only signals from the immobilized dyes are detected.

### 3. Detection of single molecules



Figures 6 shows detection of signals from single Cy3 molecule with no free dye in solution versus signals from single Cy3 molecule with background of 15nM Cy3 in solution. Fluorescence image from incorporation of Cy3 labeled nucleotide is shown in the upper panels. The signals tend to bleach in a single step, see the upper graph. When there are free labeled nucleotides in the solution (15nM free dye), the background signal is stronger (lower right panel) than the background signal in the absence of free labeled nucleotides in the solution. But the signal from the incorporated single molecule can still be detected. The ability to detect single molecule in the presence of free dye enables one to follow incorporation of nucleotide into an immobilized DNA template in real time.

The upper left panel of Figure 6 showed typical images of single molecules (see the bright spots). When the intensity of a spot is traced in real time (upper right panel), one can see that it appears (incorporation event or sticking to the surface event) and disappears (bleaching or detaching event). The same results are also illustrated in the middle long thin panel of Figure 6. This panel shows successive images of a small area around the spot that was being traced. The fluorescent signal appeared and disappeared after every few seconds (every frame is a second exposure).

**Example 5. Determining Nucleotide Incorporation Based on Correlation of Fluorescence Spots**

A correlation was observed between the position of the immobilized DNA template on the surface (indicated by the fluorescently labeled primer) and the incorporation of nucleotide to the surface. In Figure 4, image of the immobilized DNA which was hybridized to the Cy5 labeled primer was shown in the upper two panels (the middle panel is a magnified image of a small area in the left panel). The small dots in the image represent likely positions of the DNA templates immobilized on the surface. The fluorescence signals were then bleached out by a long radiation (about 1 minute) at 635nm with a 10mW laser diode. Subsequently, the polymerase and the nucleotides (including the Cy3-labeled dCTP) were added, and the mixture incubated at room temperature for about an hour. After washing, a second image of the surface was taken. This time a new set of fluorescence-labeled points appeared (see lower left two panels). The results indicate that the two sets of fluorescently-labeled points are correlated (see right panel). It is noted that the significant overlap (about 40%) between DNA primer location (Cy5) and dCTP Incorporation location (Cy3) cannot be a random result. Under the concentrations of labeled DNA primers used in the experiment, the probability for this correlation to occur randomly calculated to be about

10<sup>-50</sup>. Rather, the correlation is due to incorporation of the Cy3 labeled nucleotides into the immobilized, Cy5 labeled primer.

Incorporation of labeled nucleotide into the immobilized template is also demonstrated by the multi-incorporation data shown in Figure 7. When the intensity of the spots in Figure 4 were measured, a multistep bleaching is observed (Figure 7, upper left panel). Simulation of the multiple bleaching is shown in the upper right panel. The results are what should be expected if few molecules are located in the same place up to the optical resolution. This indicates that the polymerase can incorporate a few labeled nucleotides into the same DNA template. In a control experiment, ddATP, dCTP-Cy3 and dGTP were used to extend Cy5-labeled primer PMu50Cy5, Cy5 5' - gctggcggatgac - 3' (SEQ ID NO:4) along the Mu50 oligonucleotide template (SEQ ID NO 3). This allows only one Cy3-labeled nucleotide to be incorporated into the primer because the first codon in the template sequence after the primer is CGT. Incorporation of ddATP immediately after the incorporation of dCTP-Cy3 terminates the elongation. As shown in the lower right panel, there is no multibleaching.

It is noted that because the concentration of the DNA template on the surface was so low, it is unlikely that more than one copy of the DNA template were present on each spot. Further, multiple bleaching is not common when the polymerase was not present (data not shown). In particular, there is no correlation between primer location and fluorescence signal from the surface when the polymerase was not present (see, e.g., Figure 13, middle panel).

#### Example 6. Dynamics of Nucleotide Incorporation

Figure 8 shows a time course of incorporation events during the DNA polymerase reaction. In this experiment, the DNA template and Cy5-labeled primer complex was immobilized to the substrate surface as described above, and its position was imaged. The DNA Polymerase was then added along with the nucleotides of which one was labeled with Cy3.

As indicated in the figure, the substrate was imaged every 10 sec, with a 1 sec exposure. Every spot with immobilized DNA template (as indicated by the labeled primer) was monitored as a function of time. A series of small images of these spots were placed along a strip resulting in a movie showing the "activities" at each point.

Repeated incorporation of nucleotide into the DNA template was shown in Figure 9. Using more dyes will enable us to read the sequence of the DNA directly in an

asynchronous manner. Figure 9 shows the dynamic incorporation events at 8 different spots. The digital information recorded in these movies indicate that repeated incorporation events occurred at various time points. The data also demonstrated the feasibility of monitoring primer extension activities on single DNA molecules.

Figure 10 shows a histogram of the number of incorporation events on single spots and a histogram of the time between incorporation events. From the histograms one can see that a few nucleotides were incorporated into single DNA molecules. The low numbers of events in which more than three nucleotides were incorporated indicate that there is some mechanism that prevents high number of incorporation into the DNA under the experimental conditions. The reason could be that photo-damage to the DNA in the surrounding area of the illuminated dye might produce toxic radicals. Changing the reaction conditions and reagents could increase the numbers of incorporated nucleotides dramatically.

#### Example 7 Base-by-base Sequence Analysis

This experiment was performed to confirm selectivity of the polymerase and to illustrate feasibility of determining the sequence of a polynucleotide template with base-by-base scheme.

First, fidelity of the polymerase in incorporation was confirmed by analyzing correlation between location of immobilized primer and location of nucleotide incorporation with a correlation graph. Figure 11 shows correlation between primer location and polymerase activity location. The position of each point was determined with a sub pixel resolution. Images for the primer location and the incorporation position were taken first. If there is a correlation between the two, there is a pick in the correlation graph. Otherwise no pick was observed. As shown in the figure, the two images correlate with each other.

Results demonstrating base-by-base analysis of the sequence of a immobilized template at single molecule resolution is shown in Figure 12. The data indicated that at least two bases of the template were determined by flowing in and out reagents along with different types of labeled nucleotides (e.g., dCTP-Cy3, dUTP-Cy3, etc.). Here, a 6TA6GC oligonucleotide template (SEQ ID NO:6) was immobilized to the fused silica slide. A Cy3-labeled p7G primer (SEQ ID NO:2) was annealed to the template. As illustrated in the Figure, the primer was first extended up to the A residue with non-labeled dATP nucleotides. Then, dUTP-Cy3 nucleotide was incorporated and imaged. Images taken at this time show high correlation (see the upper left correlation graph). After bleaching the dyes, dCTP-Cy3 was applied to the sample. Images taken at this time show low correlation (see the lower left

correlation graph). Thereafter, non-labeled dGTP was added to fill the CCCCC gap till the G residue in the sequence. At this time, incorporation of a dCTP-Cy3 nucleotide was examined again. This time there was a correlation between the dCTP-cy3 positions and the primer positions in general, and in particular there was a correlation with the position of the incorporated dUTP in the first incorporation cycle. Thereafter, dUTP-Cy3 was added. Correlation was found between the labeled primer position and signal from dUPT-Cy3, but no correlation was found between the new dUPT-Cy3 positions and the position that has incorporated dUTP in the first incorporation cycle (lower right graph). The interpretation is that not all the primers were extended in the first dUTP incorporation cycle, that those which did not get extended could incorporate dUTP in the second incorporation cycle, and that those which did incorporate dUTP in the first cycle could not incorporate dUTP again in the second cycle. The results indicate that on those spots which have incorporated the first U residue there were also incorporations of a C but not a U residue. Thus, identity of a second base can be determined with the experimental scheme, although the yield for the second base (upper right graph) was not as good as for the first base (upper left graph).

In a control experiment, after filling in with A residues, dCTP-Cy3 (wrong nucleotide for the first base) was added. Correlation between Cy3-labeled primer position and C-Cy3 was low (data not shown). In another control, after filling in the string of A residues, the U residue, G residues, and U-Cy3 (wrong residue for the second base) was added. The correlation observed from the results in this experiment was low (at the noise level; data not shown). Using different oligonucleotide templates, the experiment scheme was repeated for successive incorporations of other combinations of two or more nucleotides (data not shown). The results confirmed correct incorporation of the first labeled nucleotide with high signal-to-noise ratio and subsequent incorporations of more nucleotides with a relatively lower signal-to-noise ratio. Taken together, these data indicate that the observed results (e.g., as shown in Figure 12) are not due to artifacts, but rather demonstrate efficacy of base-by-base analysis of the experimental scheme.

#### Example 8. Two Color Incorporation: Fluorescence Resonance Energy Transfer

This experiment demonstrate incorporation of two different fluorescent labels into the same immobilized polynucleotide template through detection of fluorescence resonance energy transfer (FRET). In this experiment, two fluorescent labels were used (Cy5 and Cy3), and FRET from dUTP-Cy3 (donor) to dCTP-Cy5 (acceptor) was examined at the single molecule level as shown in Figure 13.

Image of the DNA template with the labeled primer is shown in the left panel. Detection of FRET after incorporation of the two labels is provided in the right image.

Correlation between the template location and the incorporation signals is shown in the middle graph. As indicated, there is a high correlation between the template location and the incorporated nucleotide location. A control experiment was performed in which no polymerase is present. Results from the control experiment produced a low correlation between the template location and location of labeled nucleotides. FRET experiment provides particularly high signal to noise ratio as there is almost no signal from nonspecific incorporation of dyes to the surface.

When the two labels were incorporated into a primer at close vicinity, i.e., at a few nanometers apart, a single molecule FRET signal was detected (Figure 14). To detect the FRET signal, the optic setup was altered. A image splitter was added so that the same area was imaged twice (Optical Insights LTD, micro imager device). In one channel, a fluorescence filter detected only the donor (cy3) fluorescence. In the other channel, a filter for the acceptor (Cy5) was placed. With this setup individual spots were examined after incorporation. Figure 15 further indicates that the FRET detection scheme allows measurement of incorporation rate with a nice signal to noise ratio.

WHAT IS CLAIMED IS:

1. A method of analyzing sequence of a target polynucleotide, comprising:
  - (a) providing a primed target polynucleotide immobilized to a surface of a substrate; wherein the target polynucleotide is attached to the surface with single molecule resolution;
  - (b) adding a first fluorescently labeled nucleotide to the surface of the substrate under conditions whereby the first nucleotide attaches to the primer, if a complementary nucleotide is present to serve as template in the target polynucleotide;
  - (c) determining presence or absence of a fluorescence signal on the surface where the target polynucleotide is immobilized, the presence of a signal indicating that the first nucleotide was incorporated into the primer, and hence the identity of the complementary base that served as a template in the target polynucleotide; and
  - (d) repeating steps (b)-(c) with a further fluorescently labeled nucleotide, the same or different from the first nucleotide, whereby the further nucleotide attaches to the primer or a nucleotide previously incorporated into the primer.
2. The method of claim 1, wherein step (a) comprises providing a plurality of different primed target polynucleotides immobilized to different portions of the substrate.
3. The method of claim 1, wherein steps (b)-(c) are performed at least four times with four different types of labeled nucleotides.
4. The method of claim 1, wherein steps (b)-(c) are performed until the identity of each base in the target polynucleotide has been identified.
5. The method of claim 1, further comprising an additional step of removing the signal after step (c).
6. The method of claim 1, wherein the presence or absence of a fluorescence signal is determined with total internal reflection fluorescence (TIRF) microscopy.

7. The method of claim 1, wherein the target polynucleotide is primed with a fluorescently labeled primer.
8. The method of claim 1, wherein the first and further nucleotide are labeled with the same fluorescent label.
9. The method of claim 1, wherein said the substrate is a fused silica slide.
10. The method of claim 9, wherein said surface is coated with a polyelectrolyte multilayer (PEM).
11. The method of claim 10, wherein said PEM is terminated with a polyanion.
12. The method of claim 11, wherein said polyanion bears pendant carboxylic acid groups.
13. The method of claim 12, wherein said target polynucleotide is biotinylated, and said surface is coated with streptavidin.
14. The method of claim 13, wherein said surface is coated with biotin prior to coating with streptavidin.
15. The method of claim 14, wherein said surface is coated with a polyelectrolyte multilayer (PEM) terminated with carboxylic acid groups prior to attachment of biotin.
16. The method of claim 1, wherein said removing or reducing is by photobleaching.
17. The method of claim 1, wherein the substrate is in fluid communication with a microfluidic device, wherein the first and further labeled nucleotides are added to or removed from the substrate through the microfluidic device.
18. The method of claim 17, wherein the microfluidic device comprises  
(a) a flow cell comprising the substrate; and

(b) an inlet port and an outlet port, said inlet port and outlet port being in fluid communication with said flow cell for flowing fluids into and through said flow cell.

19. The method of claim 18, wherein the substrate is a microfabricated synthesis channel.

20. The method of claim 17, further comprising a light source to illuminate the surface of said substrate and a detection system to detect a signal from said surface.

21. The method of claim 17, further comprising an appropriately programmed computer for recording identity of a nucleotide when said nucleotide becomes incorporated into the target polynucleotide.

22. A method of analyzing sequence of a target polynucleotide, comprising:

(a) providing a primed target polynucleotide immobilized to a surface of a substrate; wherein the target polynucleotide is attached to the surface with single molecule resolution;

(b) adding four types of nucleotides to the surface of the substrate under conditions whereby nucleotides attach to the primer dynamically, when complementary nucleotides are present in the target polynucleotide; and

(c) monitoring in a time course of incorporation of fluorescent signals into the immobilized primer.

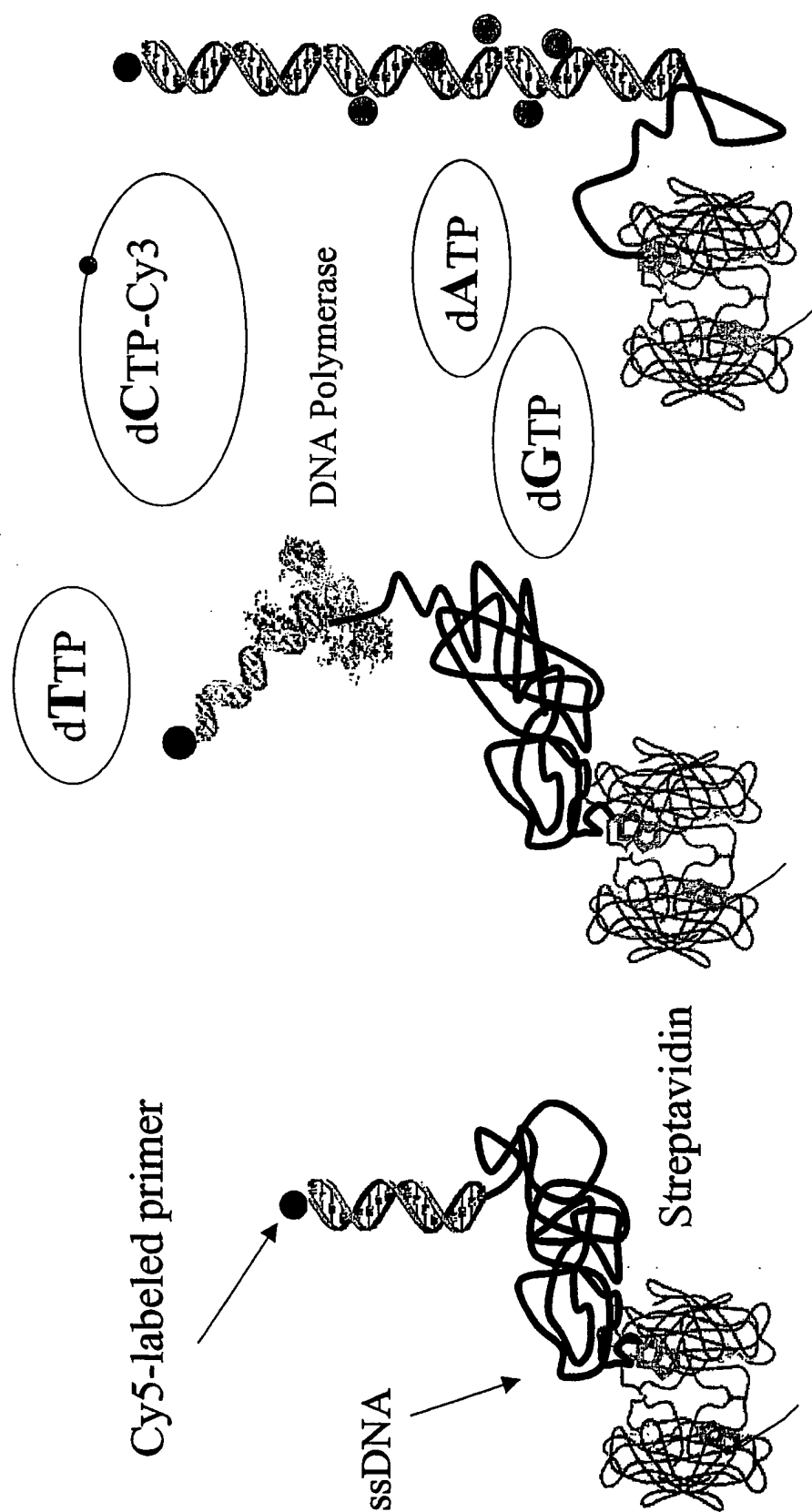
23. The method of claim 22, wherein monitoring of fluorescent signal incorporation into the immobilized primer is by taking images in a time course with monitored with total internal reflection fluorescence microscopy.

24. The method of claim 23, wherein the images are taken at a rate faster than the rate at which nucleotides are incorporated into the primer.

25. The method of claim 23, wherein nucleotide concentrations are low at each time point when an image is taken.



26. The method of claim 25, wherein nucleotide concentrations are alternated by fluid exchange with a microfluidic device.
27. The method of claim 22, wherein all four types of nucleotides are each labeled with a different label.
28. An apparatus for analyzing the sequence of a target polynucleotide, comprising:
- (a) a flow cell comprising a substrate for immobilizing the target polynucleotide with single molecule resolution;
  - (b) an inlet port and an outlet port, said inlet port and outlet port being in fluid communication with said flow cell for flowing fluids into and through said flow cell;
  - (c) a light source for illuminating the surface of the substrate; and
  - (d) a detection system for detecting a signal from said surface.



**FIG. 1**

2/15

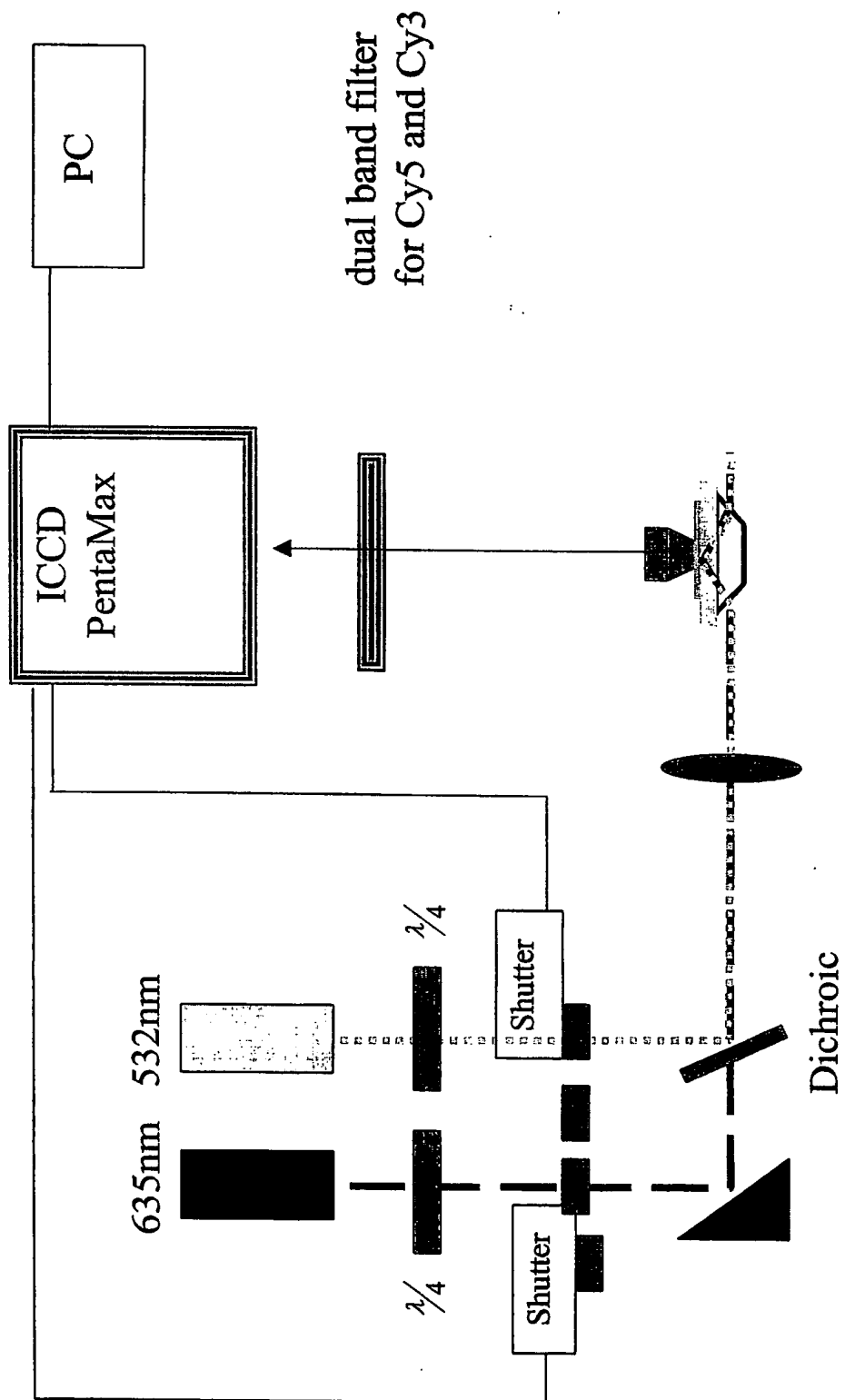
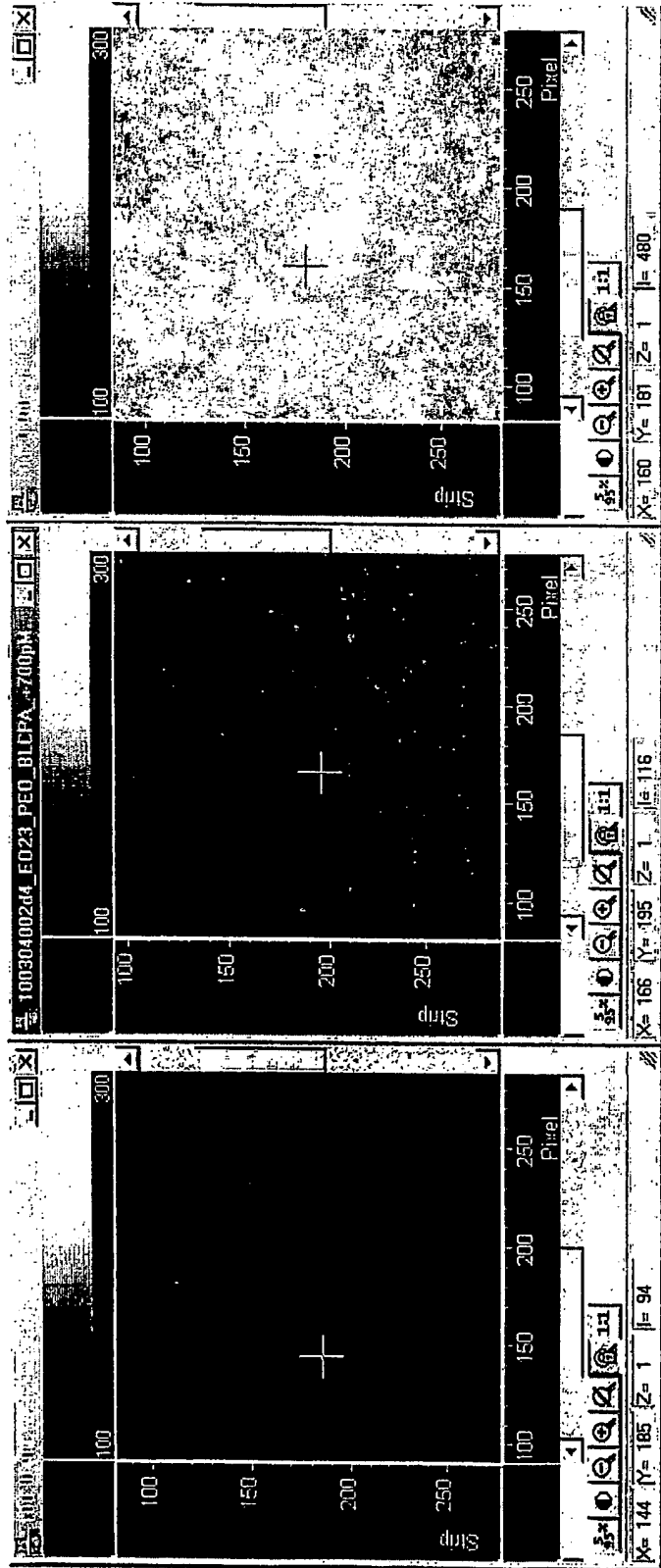


FIG. 2

3/15



Streptavidin Without dye

0.7nM Cy5-primer+ Biotin\_Oligo, Without Streptavidin on the surface

0.7nM Cy5-primer+ Biotin\_Oligo, With Streptavidin

FIG. 3

4/15

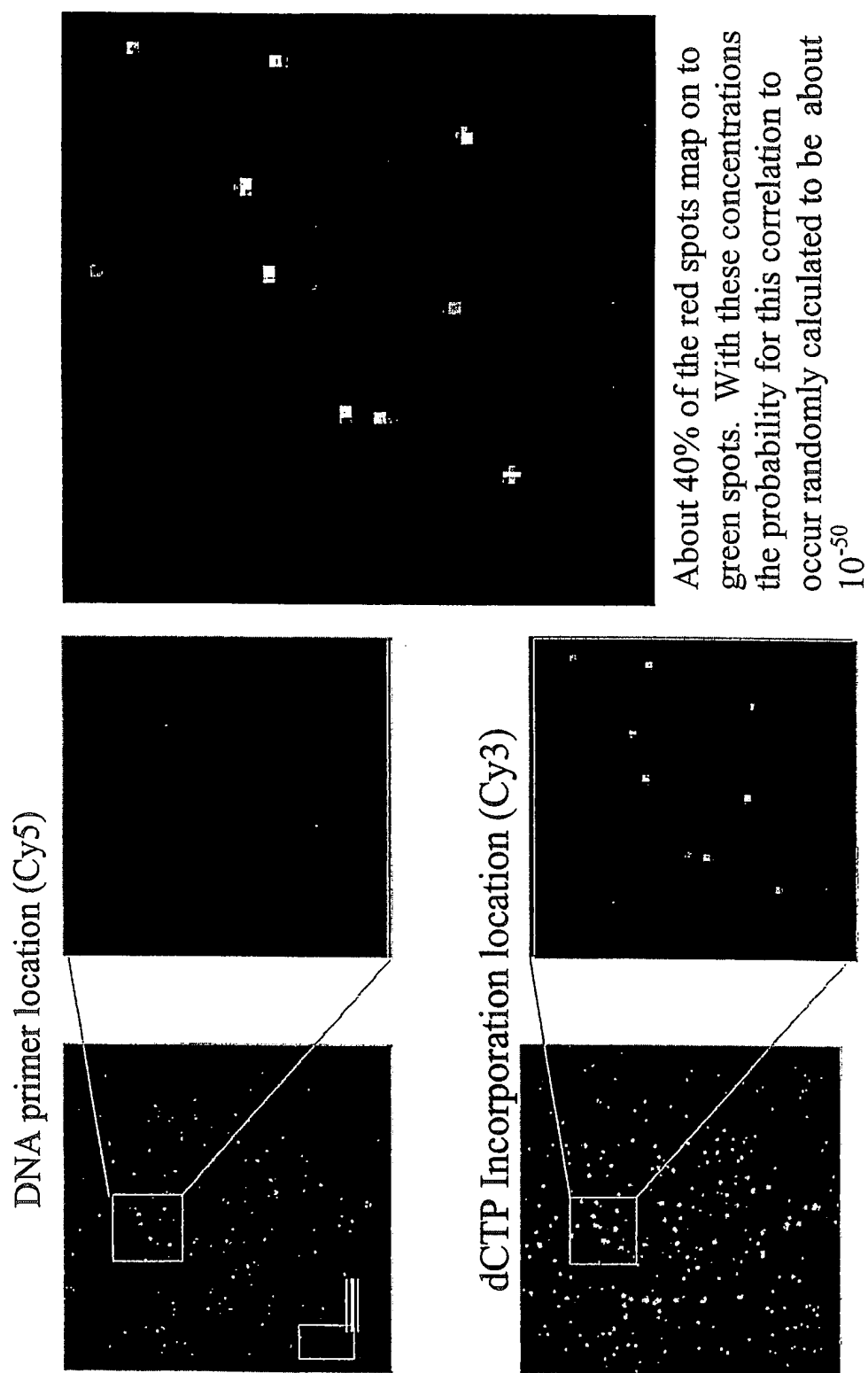


FIG. 4

5/15

Biotin-5'-TCAGTCATCAGTCATCAGTCATCAGTCATCAGTCATCAGTCATCAG-  
 12 11 10 9 8 7 6 5  
 -TCATCAGTCATCAGTCATCAGTCATCAGTCATCAGTCATCAGTCATCAGTTCTA-3'  
 4 3 2 1 0 3'-TGTGCCTCCAAGAT-5'-

Cy3-dCTP:ddCTP 3:1

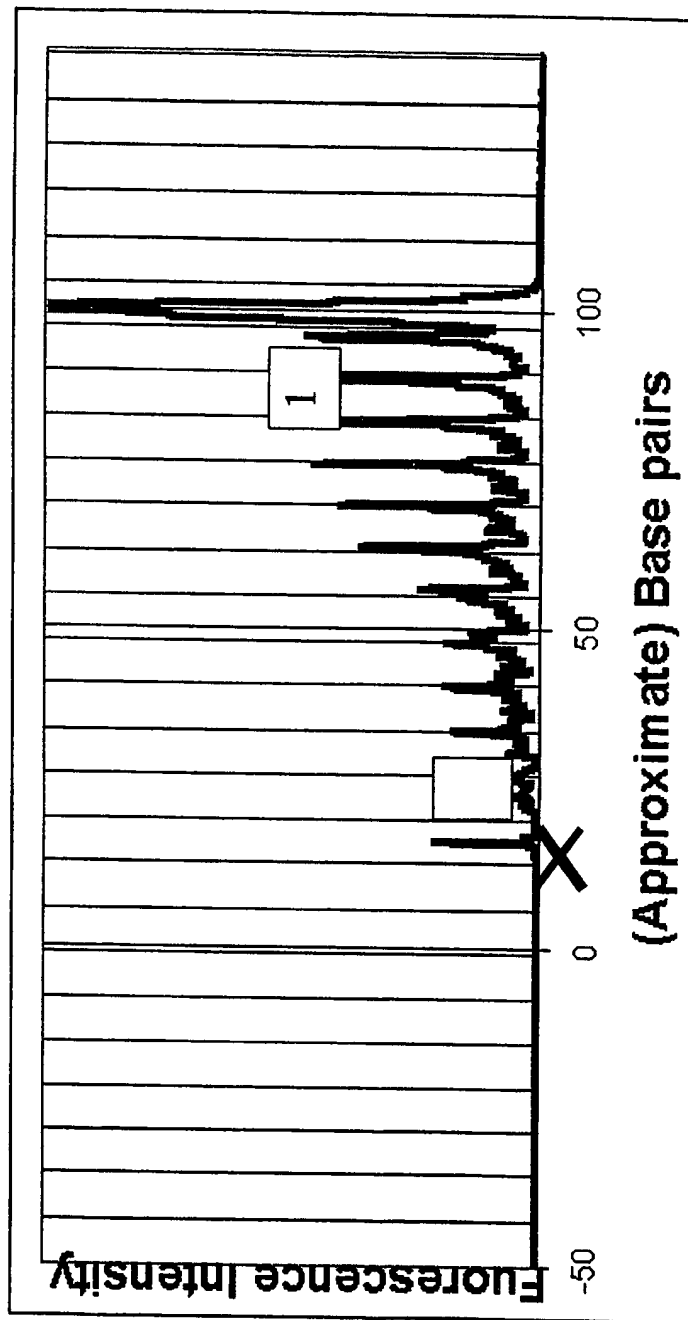


FIG. 5

6/16

- clean controls
- bleaching in one step
- number of objects correlates with with concentration of fluorophores
- histogram has only one peak

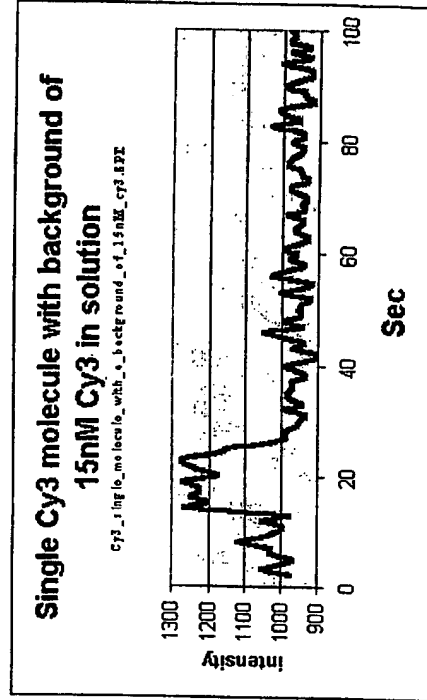
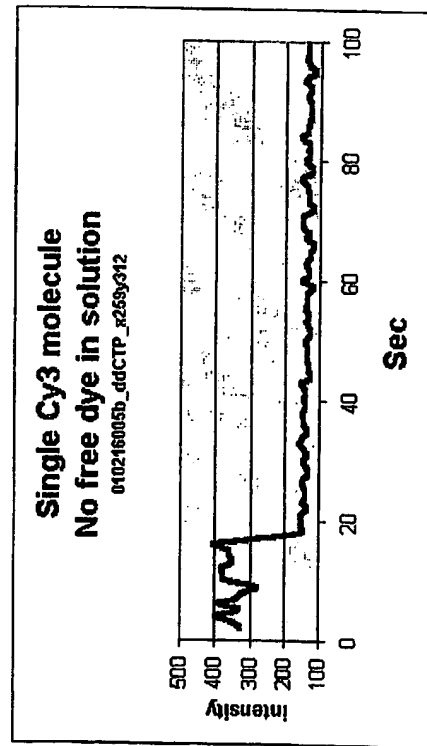
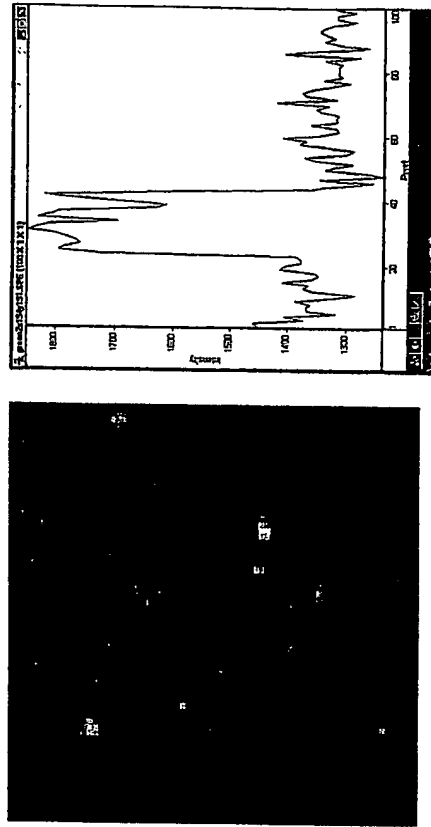


FIG. 6

7/15

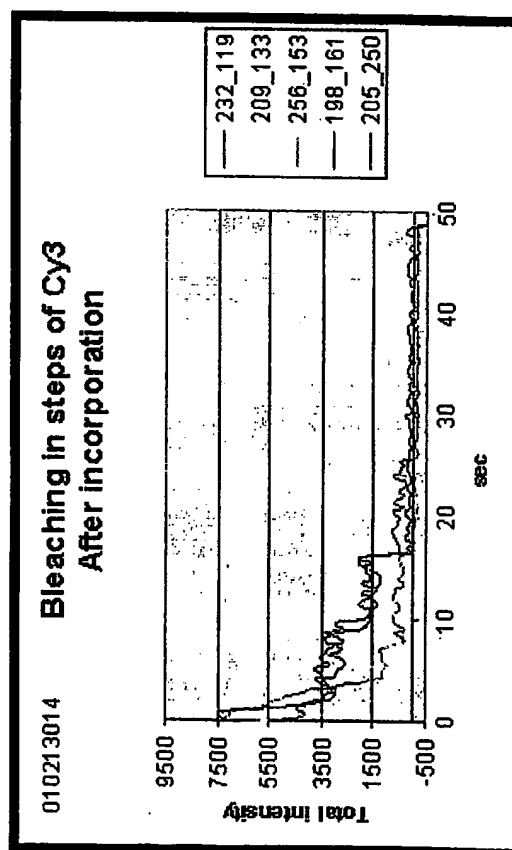
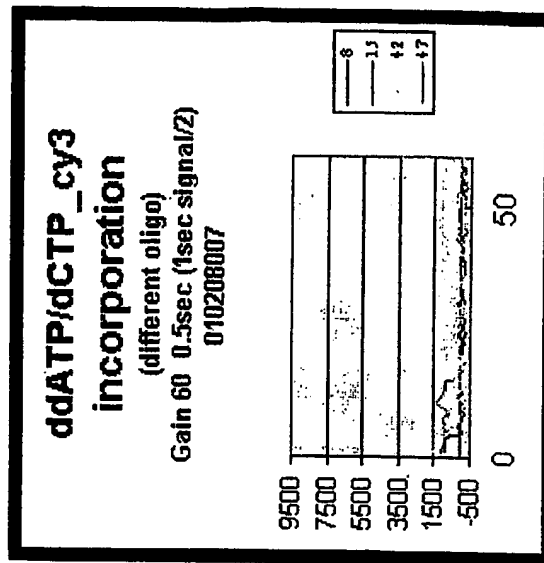
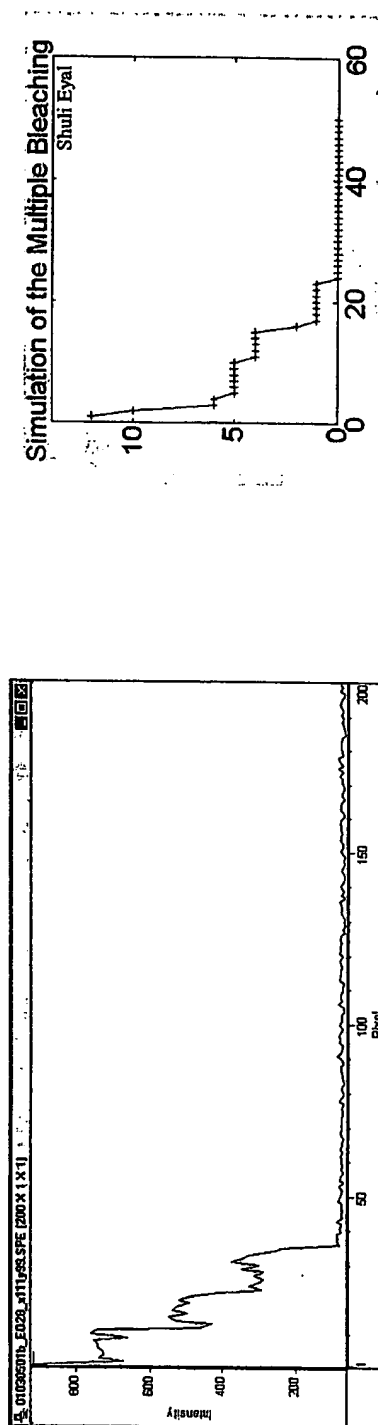
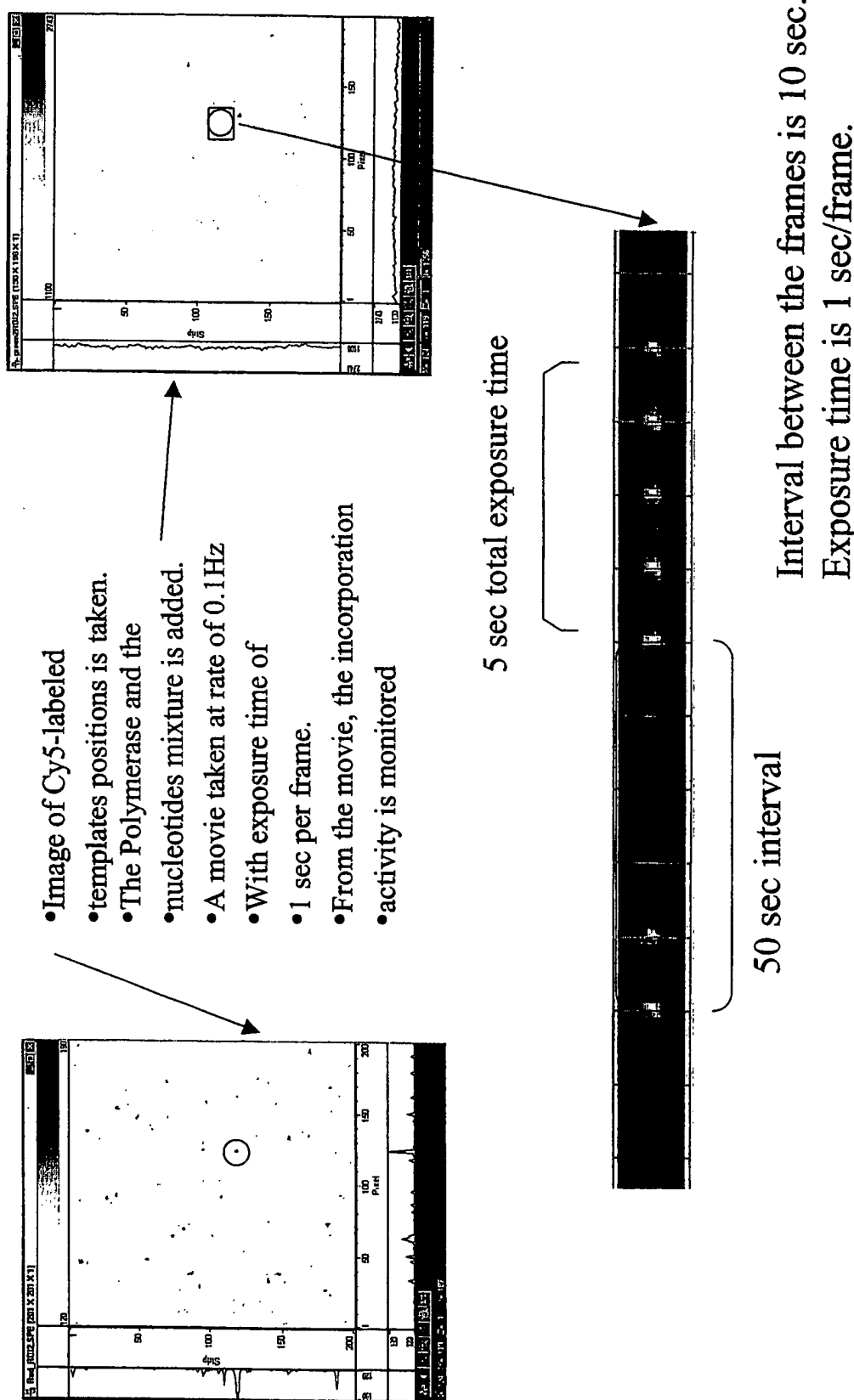


FIG. 7



8/15



9/15

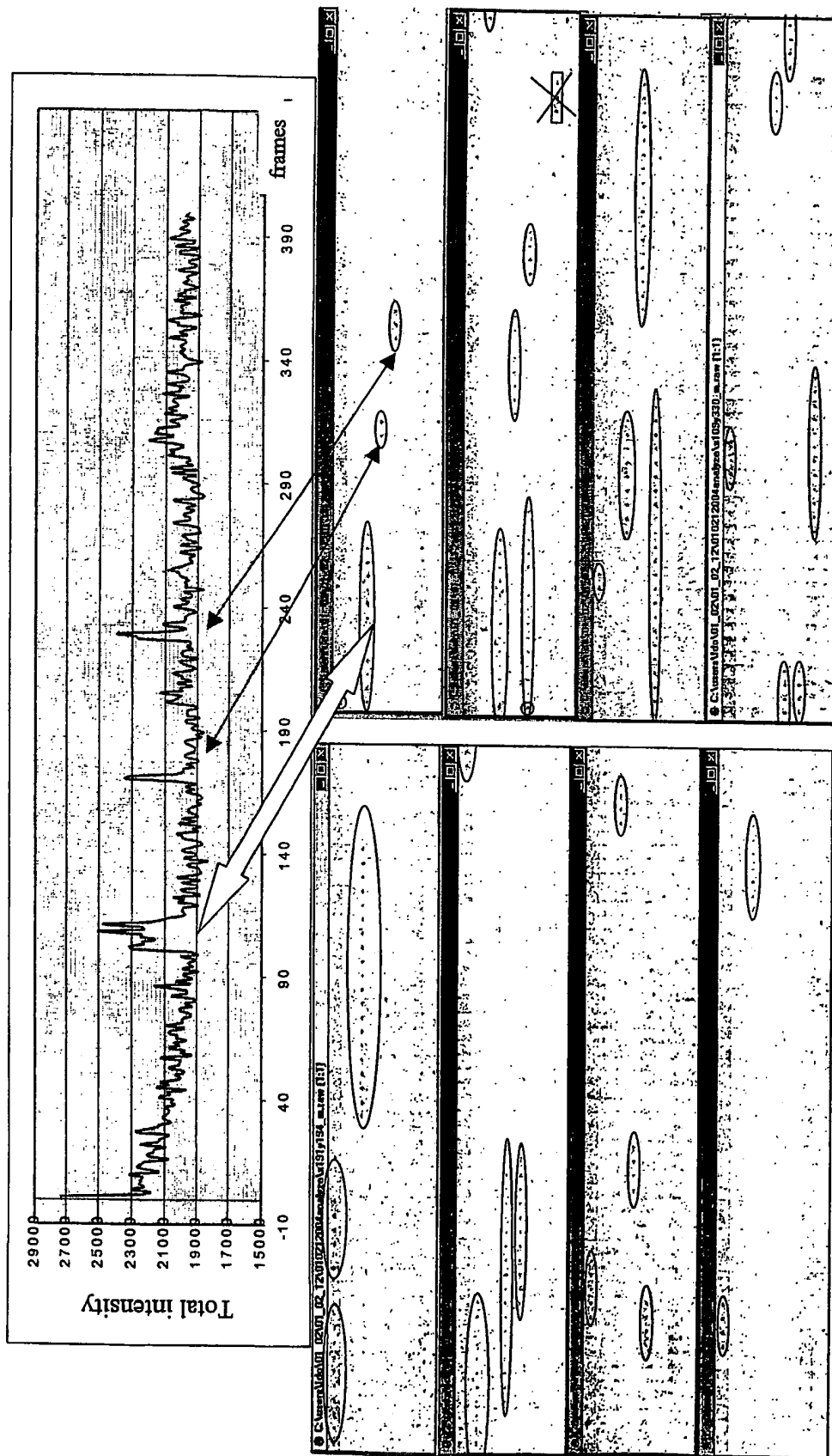


FIG. 9

10/15

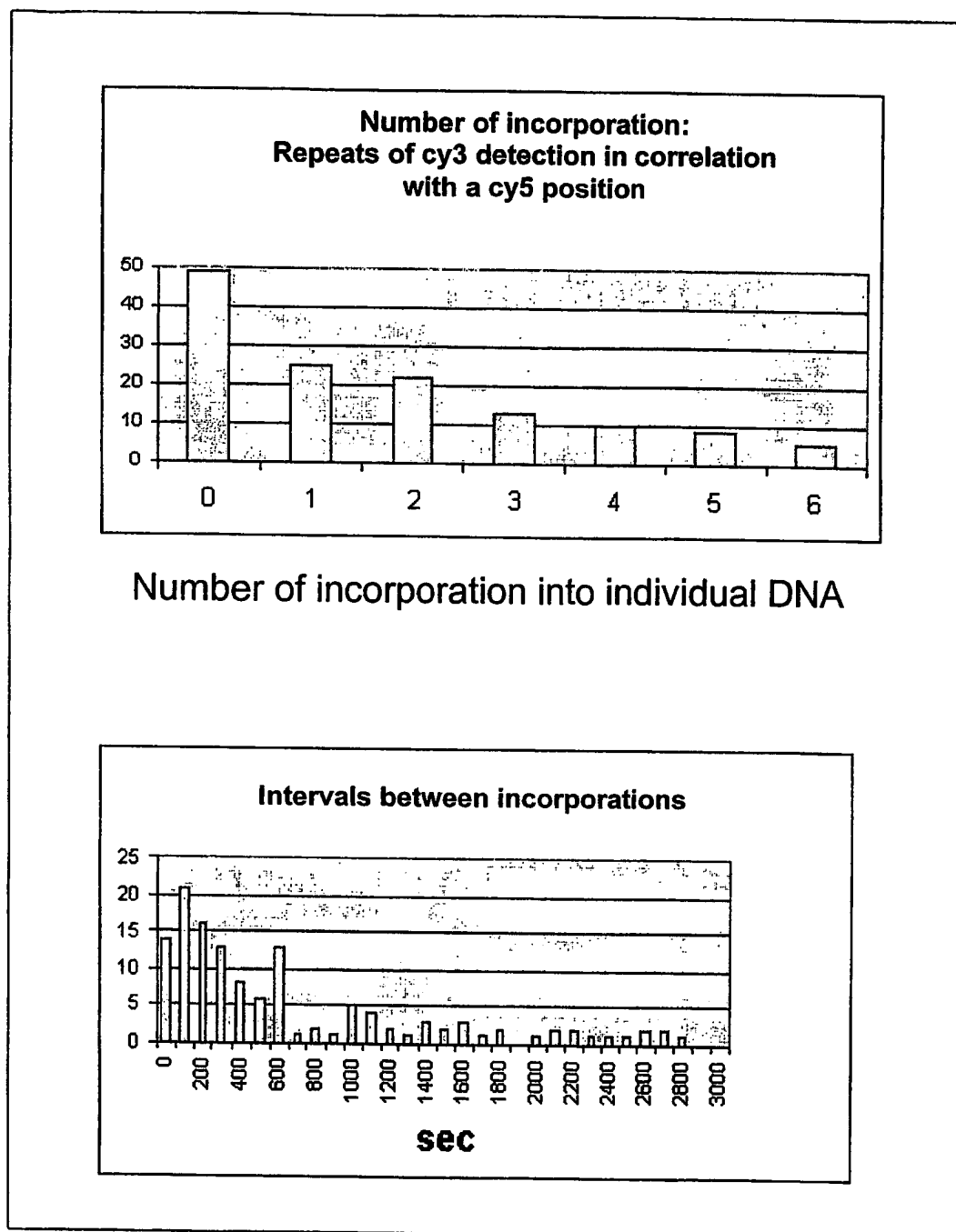
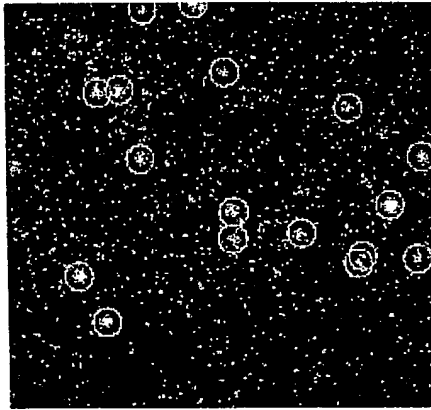


FIG. 10

11/15



Incorporation features

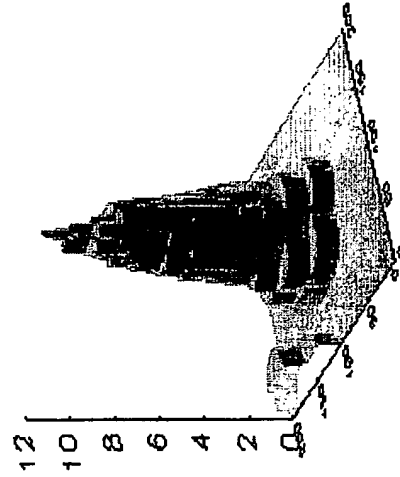
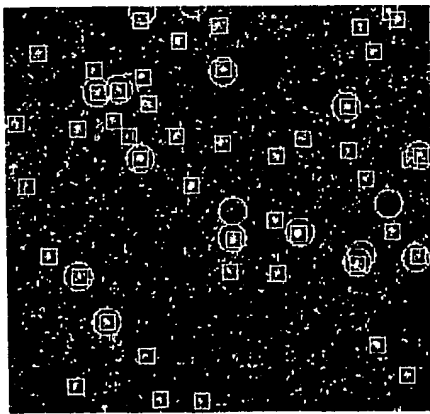
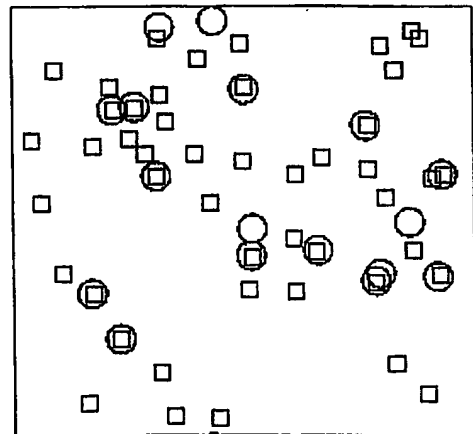


FIG. 11



primer and incorporation features



12/15

DNA Sequence: 3'... ttttttA c c c c c c G 5'  
 aaaaaaU g g g g g g C

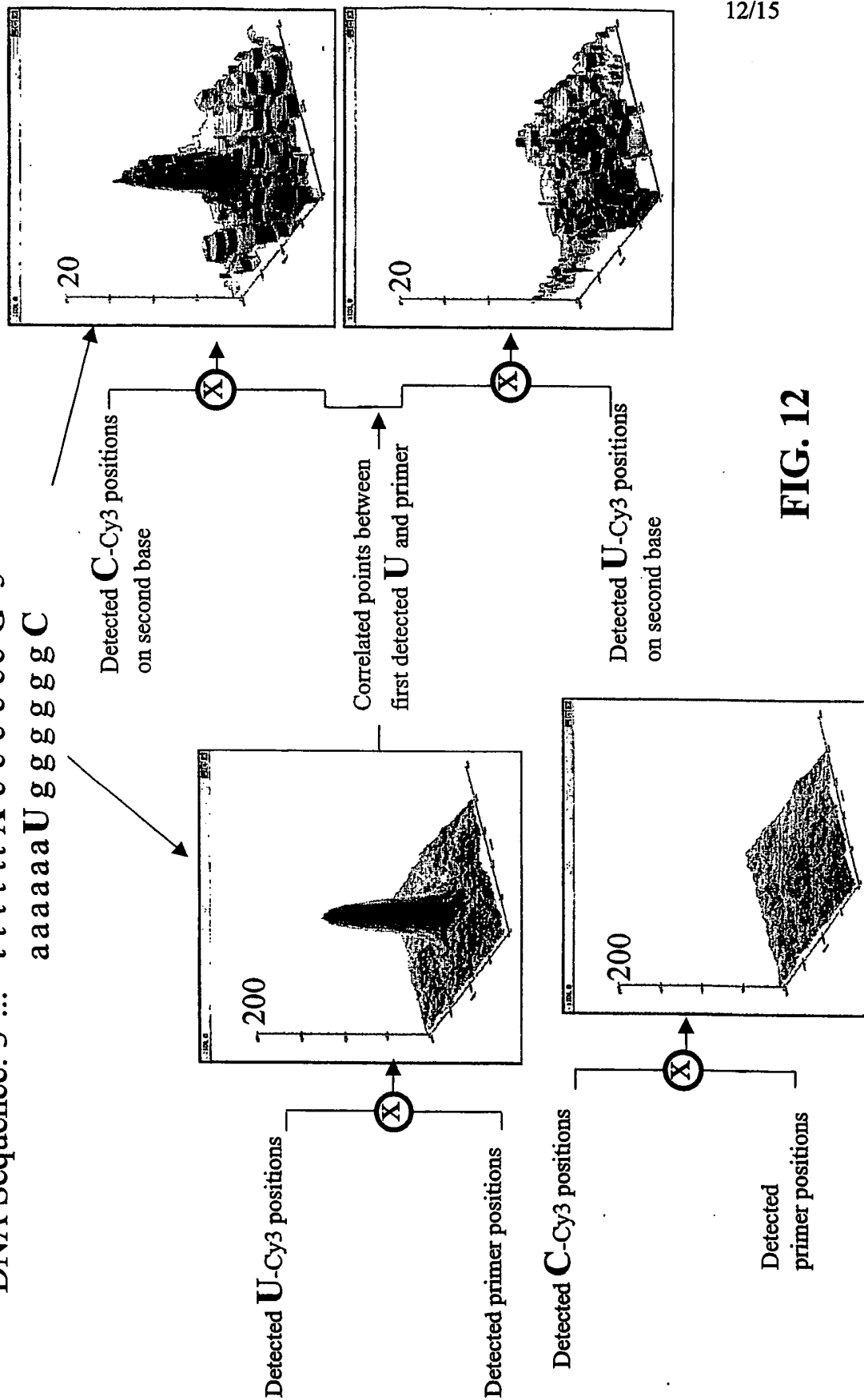
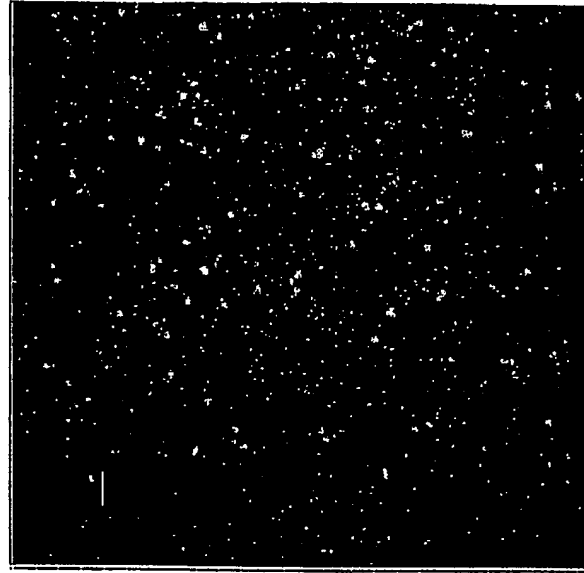


FIG. 12

13/15

•FRET signal after incorporation labeled nucleotide



•Fluorescently labeled DNA



...AttctttGcttcttAttctttGcttcttAttctttG....

FIG. 13

14/15

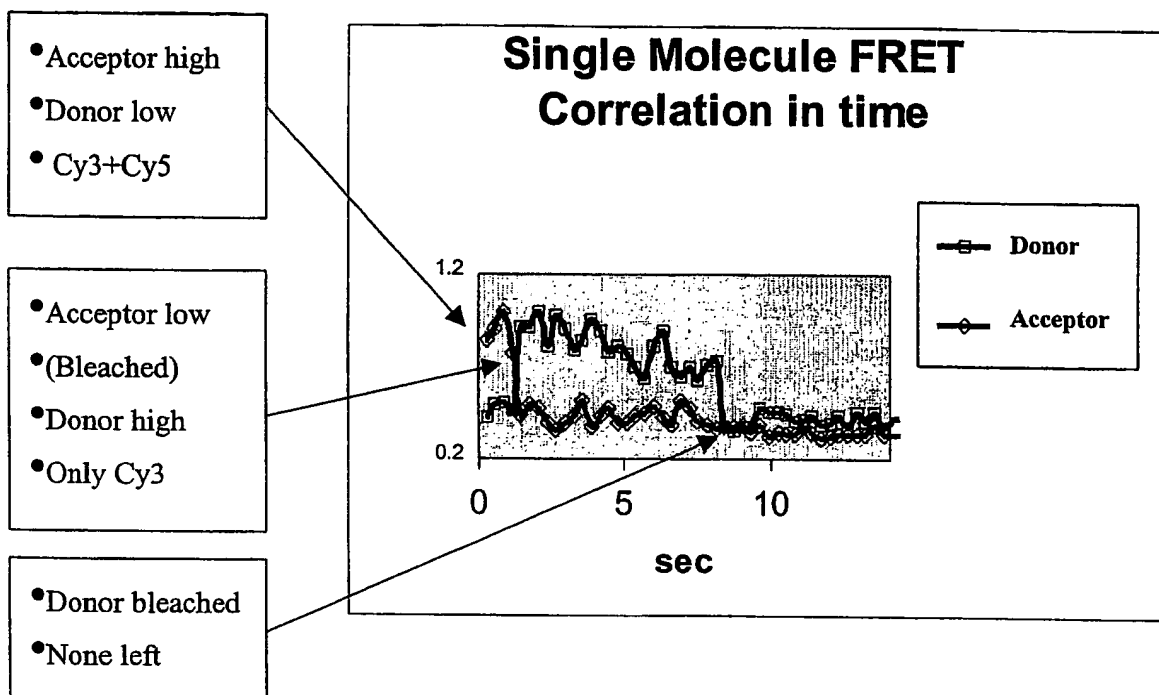


FIG. 14

DNA Sequence: 3' ...tttttttA ccccccG...5'  
 ...aaaaaaU ggggggC



Desired signal for measuring the incorporation

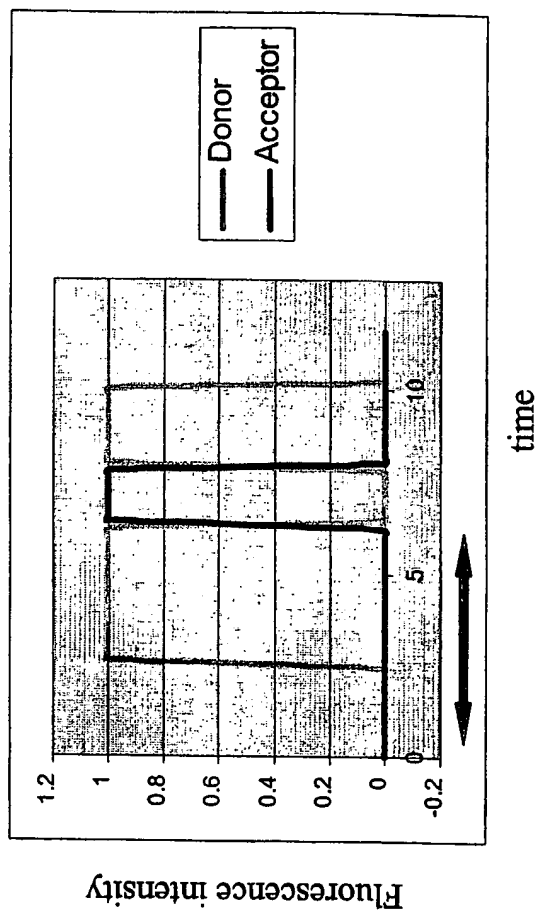


FIG. 15



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/08187

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/68; C07H 21/02, 21/04

US CL : 435/6; 536/23.1, 24.3

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
U.S. : 435/6; 536/23.1, 24.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
Please See Continuation Sheet

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,302,509 A (CHEBSEMAN) 12 April 1994, see entire document.	1, 3-5, 22, 27
Y		2, 6-21, 23-26
Y, P	US 6,255,083 B1 (WILLIAMS) 03 July 2001, see entire document.	6, 17-19, 21 and 26
Y	US 5,832,165 A (REICHERT et al) 03 November 1998, see entire document.	6, 17-19, 21 and 26
Y	US 6,020,457 A (KLIMASH et al) 01 February 2000, see entire document.	9-15
A,P	US 6,221,592 B1 (SCHWARTZ et al) 24 April 2001, see entire document.	1-27

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"Z"

document member of the same patent family

Date of the actual completion of the international search

26 June 2002 (26.06.2002)

Date of mailing of the international search report

09 AUG 2002

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks

Box PCT

Washington, D.C. 20231

Facsimile No. (703)305-3230

Authorized officer

Eden Whisenant, Ph.D.

Telephone No. (703) 308-0196

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/08187

## BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claim(s) 1-27, drawn to a method of analyzing the sequence of a target ppolynucleotide.

Group II, claim(s) 28, drawn to an apparatus for analyzing the sequence of a target ppolynucleotide.

The inventions listed as Groups I and II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The claims as drawn are related to each other because they both relate to analyzing the sequence of a target polynucleotide. However, since the method of analyzing the sequence of a target polynucleotide, as claimed, is known, see for example, Cheeseman [US 5,302,509 (1994)], the claims are no longer linked by a special technical feature, because, by definition, the a special technical feature must distinguish over the prior art. Without the special technical feature the claims lack unity.

**Continuation of B. FIELDS SEARCHED Item 3:**  
USPATFULL via EAST, Medline, CApus

search terms : sequencing, primer extension, polynucleotide or nucleic acid, fluor? and (TIRF ot total internal reflection fluorescence)

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/08187

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claim Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claim Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:  
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-27

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER: \_\_\_\_\_**

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**